THEORY OF PHYSICS

BY

Sweetman

JOSEPH S. AMES, Ph.D.

ASSOCIATE PROFESSOR OF PHYSICS AND
SUB-DIRECTOR OF THE PHYSICAL LABORATORY IN
JOHNS HOPKINS UNIVERSITY

NEW YORK AND LONDON
HARPER & BROTHERS PUBLISHERS
1897

PREFACE

To present successfully the subject of Physics to a class of students, three things seem to me necessary: a text-book; a course of experimental demonstrations and lectures, accompanied by recitations; and a series of laboratory experiments, mainly quantitative, to be performed by the students themselves under the direction of instructors. I place "text-book" first, because for many reasons I believe it to be the most important of the three. None but advanced students can be trusted to take accurate and sufficient notes of lectures; and a text-book which states the theory of the subject in a clear and logical manner so that recitations can be held on it, seems to me absolutely essential. Another great advantage of a text-book is that a student provided with one does not need to take notes on the lectures, and so can give his undivided attention to the explanations and demonstrations which are being presented.

It has been my aim in writing this book to give a concise statement of the experimental facts on which the science of Physics is based, and to present with these statements the accepted theories which correlate or "explain" them. I have in no case resorted knowingly to an antiquated or incorrect theory for the sake

of ease of demonstration; and I have constantly tried to show the connections between different phenomena, and to reduce the general laws of nature to as small a number as possible. Comparatively little attention is given in the text to the description or discussion of the great fundamental experiments of Physics; because this can be given in the class demonstrations; and, in any case, this side of the subject deserves consideration in a separate book. The central thought of the present book is the *theory* of the experiments, and their "explanation" in terms of more fundamental ideas and principles. It is for this reason that I have chosen the title "Theory of Physics." I am not at all certain that the choice is a happy one; but it at least has the merit of emphasizing what I have desired to make the chief characteristic of the book.

As any teacher knows, it is well-nigh impossible to give courses of lectures on any subject without from time to time adopting modes of expression and explanation, even outlines of certain theories, from various books and articles; and I am perfectly conscious, in my own case, of having received great assistance in numberless ways from many writers. The present book is in the main simply a revision of a course of lectures which I have been giving for some years; and I am certain that many ideas which are not my own have crept in. They are now, however, so assimilated, even in my mind, that it is absolutely impossible for me to distinguish that which is strictly original from that which is borrowed. If any one wishes to claim any portion of the book as his and not mine, I shall be the first to try to

recognize his claim; but I can give the strictest assurance that in no case have I deliberately used even a sentence or a single idea which belongs to another.

The present text-book is designed for those students who have had no previous training in Physics, or at the most only an elementary course; and it should, then, be adapted to junior classes in colleges or technical schools. The entire subject as here presented may be easily studied in a course lasting for the academic year of nine months. Although there is a close connection between the various sections, I have tried to divide the subject in such a way that it need not all be considered in one continuous course; for in many colleges only selected portions, e. g. Heat and Light, are taken up at a time. I ought, perhaps, to state the reasons which have led me to make the divisions and to arrange them as I have; but in reality I have but one reason: I have constantly had in mind the needs of the student, and have tried in every way to treat the subject so as to make it clearest to him.

I wish to express my sense of obligation to my assistant, Mr. N. E. Dorsey, who has kindly read my entire manuscript and called my attention to many points demanding correction; and also to Mr. C. L. Reeder, who has made all the drawings most skilfully and promptly.

J. S. AMES.

JOHNS HOPKINS UNIVERSITY,
BALTIMORE, October, 1896.

CONTENTS

Introduction

BOOK I

MECHANICS AND PROPERTIES OF MATTER

CHAPTER I

KINEMATICS

CHAPTER II

DYNAMICS

CHAPTER III

GRAVITATION

CHAPTER IV

PROPERTIES OF SIZE AND SHAPE OF MATTER

BOOK II

SOUND

CHAPTER V

VIBRATING BODIES

CHAPTER VI

VELOCITY OF SOUND

CHAPTER VII

HARMONY AND MUSIC

BOOK III

HEAT

Introduction

CHAPTER I

HEAT-EFFECTS — TEMPERATURE

CHAPTER II

CHANGES IN VOLUME

CHAPTER III

CHANGE IN TEMPERATURE

CHAPTER IV

MOLECULAR CHANGES

CHAPTER I

GENERAL PROPERTIES OF ELECTRIC CHARGES

CHAPTER II

ELECTRIC POTENTIAL AND INDUCTION

CHAPTER III

ELECTRIC CONDUCTION

CHAPTER IV

PROPERTIES OF STEADY ELECTRIC CURRENTS

CHAPTER V

GENERAL PROPERTIES OF MAGNETS AND MAGNETIC FIELDS

CHAPTER VI

MAGNETISM OF THE EARTH

CHAPTER VII

MAGNETIC PROPERTIES OF STEADY ELECTRIC CURRENTS

CHAPTER VIII

INDUCED CURRENTS

CHAPTER IX

GENERAL PROPERTIES OF A MAGNETIC FIELD

BOOK V

LIGHT

INTRODUCTION

CHAPTER I

THE WAVE THEORY

CHAPTER II

REFLECTION

CHAPTER III

REFRACTION

CHAPTER IV

DISPERSION — SPECTRA

CHAPTER V

COLORS — COLOR-SENSATION

CHAPTER VI

INTERFERENCE — DIFFRACTION

CHAPTER VII

DOUBLE REFRACTION

CHAPTER VIII

POLARIZATION

THEORY OF PHYSICS

INTRODUCTION

1. Physics. As we observe the world of nature around us, we are constantly receiving sensations caused by phenomena which are independent of ourselves. Often we are able to determine the exact conditions which produce a certain phenomenon; and, as the result of observations which have been carried on for many hundred years, it may be stated that there is every evidence for believing that, if the exact conditions are ever reproduced, the same result always follows. This is equivalent to saying that there are certain laws in nature which are entirely independent of time, of place, and of men. It is the aim of Natural Science to discover these laws. Physics, a branch of natural science, is concerned with only a limited number of these phenomena of nature and the laws according to which they take place; but it is practically impossible to give a definite limitation to its field. Physics includes the study of many properties of matter, such as its inertia, its weight, and its form; also the consideration of the phenomena of sound, of heat, of light, of electricity and magnetism. Physics, of course, comes closely in contact with many other natural sciences, especially with Chemistry and Astronomy.

2. Matter. The name "matter" is ordinarily given to all those substances which in any way appeal to the senses. This must not be regarded as a definition; for that can be given only in a description of the properties of matter. The word "matter" conveys a fairly definite idea to our minds simply as the consequence of our daily experiences; but if we analyze our ideas of it, we see that among them all there are three which are apparently independent of each other, and which in this sense are fundamental. Our original consciousness, though, of these fundamental properties of matter depends upon only one of our senses, that of muscular resistance.

If we in any way alter the motion of a piece of matter, we are conscious of a certain muscular exertion. Of course this same change in the motion of the matter could be produced by other causes than our muscles; but our own conception of matter depends on our muscular sense. Illustrations of this property are afforded when we throw or stop a ball, when we open a door or push a barrel along a floor; and it has received the name "inertia."

If we raise a piece of matter from the floor to the top of a table, — that is, if we move a portion of matter further from the earth, — we are conscious of muscular exertion. And, so far as we can see, there is no connection between this fact and the property of inertia. We shall learn later that this consciousness of muscular effort when a piece of matter is raised from the earth is only a special case of a much more general property; for there seems to be reason for believing that if we were to increase the distance apart of *any two portions* of matter, however small or large, we should be conscious of muscular sensation, if only our senses could be delicate enough. This property of matter is called its "weight."

If we alter in any way the shape or volume of a portion of matter, we are, in general, conscious of muscular exertion. Illustrations of this fact are given when we stretch

or twist a wire, squeeze out thin a piece of putty, compress a gas in a rubber bag, etc. Various names have been given to this property, depending upon the particular effect produced. These will be discussed later under the names "elasticity," "plasticity," "ductility," etc.

Any substance having these three properties is called matter. There is no reason for thinking that they are all independent of each other; but at present we have no evidence which will allow us to regard one property as a consequence of the others, except in the case of gases, where we can show that the elasticity is a direct result of inertia.

It must be carefully borne in mind that, although the ideas of the properties of matter come entirely through the ordinary senses, we may have conclusive evidence as to the nature of other things which in no way appeal to the senses. We shall have several illustrations of this in the course of our study.

3. Forms of Matter. That different kinds of matter appear differently to our senses is evident to all. Perhaps the most obvious fact is that they may have different volumes and shapes; and it is by means of these properties that we ordinarily distinguish between them. The names "solid," "liquid," and "gas" have been given to certain forms of matter, whose properties differ widely.

We think of a "solid" as having a definite size and shape which are independent of its position or condition, in general. Thus a stone may rest on a table or in one's hand, or it may be falling through the air; it will have the same volume and the same shape.

A "liquid," however, although it has a definite volume, assumes the shape of the vessel which contains it. If water is poured out of one pitcher into another of different shape, it will occupy the same volume in the two vessels, but the shape will be entirely changed. Unless the liquid completely fills the containing vessel, there will always be

a surface of separation between the liquid and the air or whatever gas fills the rest of the vessel. This surface is called the "free surface," and has some most interesting properties, one of which is that it enables liquids to form "drops" when they escape from the containing vessel and fall towards the earth.

A "gas" if placed inside a closed vessel diffuses uniformly throughout the entire space; so that its volume as well as its shape depends solely on the vessel, not on the gas itself. Most gases are transparent, and so cannot be seen; but their properties may be studied in other ways.

As we shall see later on, these states of matter — solid, liquid, and gas — are only three particular ones out of a great many. There is an indefinite number of intermediate states which cannot be called solid, liquid, or gaseous. Gases and liquids are sometimes called "fluids," because they possess in common the property of flowing; that is, they assume the shape of the containing vessel.

All these forms of matter can be proved to have inertia, weight, and elasticity; although this fact is not so evident at first sight in the case of gases. This is simply owing to the difficulty of detecting small effects by means of our senses. If large amounts of a gas are taken, or if instruments more delicate than our senses are used, the ordinary properties of matter may be easily shown to hold for gases as well as for solids and liquids.

4. Divisibility of Matter. That there are spaces between the portions of matter making up a substance is shown by many facts of daily experience. This is self-evident in the case of wood, cork, sandstone, etc., and in all gases. Also, gases can in many cases pass through solids, as carbon-dioxide through red-hot iron. If ordinary salt is thrown in water, it is evidently broken up into parts which are so small that they can pass between the small portions of the water; for the bulk of the water and salt when mixed is not the sum of the volumes of the two separately.

We can imagine, then, a portion of matter divided into parts, and these in turn into still finer parts; and so on, each of the parts having the same properties as the original substance. Thus a piece of chalk can be broken into two pieces of chalk, each of these into two other pieces of chalk; and so on. A limit, though, will be reached when a piece of chalk is obtained which is so small that if it is separated into portions, these cease to have the properties of chalk. This final piece of chalk is called a "molecule" of chalk. So molecules of other substances can be defined as those portions of the substances which are so small that if they are separated into parts, these will have entirely different properties. Any piece of matter can, then, be regarded as made up of molecules; and experiments show that if the matter is homogeneous, the molecules are identical in every respect. These molecules are, in general, groups of still smaller portions, which, so far as we now know, cannot be subdivided. In certain substances not alone are the molecules identical, but the molecular subdivisions also; and such a body is said to be "elementary" or an "element." Thus hydrogen is an element; because not alone are its molecules alike, but the ultimate portions of the molecule are, so far as we now know, identical. Similarly oxygen is an element. Steam is not an element, because, although the molecules are identical, the portions of a molecule are not alike, but consist of two portions like the final portions of hydrogen and one portion like the final portions of oxygen. Thus steam is said to be a "compound" of hydrogen and oxygen. Of course, the word "final" is used simply to mark the limit reached at present; further research may produce still further subdivision. But this branch of science which considers the divisions of molecules and their rearrangement into compounds, belongs to Chemistry. For the purposes of Physics it is sufficient to regard all matter as groups of molecules, the molecules of any pure substance being identical in all

respects; and to remember that portions of a molecule have properties which are different from those of the molecule itself.

5. Kinetic Theory of Matter. As will be proved later on, there is conclusive evidence that all the smallest portions of matter are in motion. The material body itself may not move; but the molecules and the portions of the molecules are in constant motion. The differences between solids, liquids, and gases depend largely on the freedom which the molecules have for motion. In a gas the molecules can move with almost perfect freedom. They cannot move very far without meeting other molecules or the walls of the vessel, when their motion is changed; but, owing to the comparatively large spaces between molecules, there is great freedom. In a liquid, molecules can also move from one portion to another with comparative freedom; but the molecules are here so close together that they influence each other greatly. In a solid the molecules cannot move from one point to a distant one. The molecules of a solid may best be regarded as forming a framework or stable configuration, which is constantly in vibration, owing to the vibrations of the molecules themselves; that is, the molecules of a solid do not leave a definite position, but simply vibrate around it or through it. The passage, then, of matter from the state of a solid into that of a liquid and then into that of a gas consists in its molecules being given greater and greater freedom of motion. This view of the properties of matter is often called the "Kinetic Theory."

6. Quantity of Matter. To compare different amounts of matter, any one of the general properties of matter may be taken as the basis of comparison. Thus, if two portions of matter have the same inertia, or if they have the same weight with reference to the earth at a certain point, they may be *defined* as having the same quantities of matter. The quantity of matter in a body is ordinarily called its

"mass;" and so equal masses may be defined in either of these two ways. As will be seen later on, if two bodies have the same inertia, they also have the same weight; so that both modes of measurement of masses will give the same results. The actual methods of measurement will be described in full further on.

If any number of substances are placed inside a vessel so made that no matter can pass in or out, and if these substances react on each other in any way, or if there are changes of any kind, it may be proved by careful experiments that the mass is the same at all times. This fact is sometimes called the "Principle of the Conservation of Matter," and it further asserts that the amount of matter in the Universe is constant. The science of Chemistry is based upon this principle.

7. Motion. Since, then in any natural phenomenon the amount of matter is not altered, the only change possible, if matter is involved, is a change in the position of the matter. The body may move as a whole, or its smaller portions may move; but in all cases every phenomenon of matter must be reduced to a question of motion. *So to "explain" every phenomenon is the aim of scientific research.*

Motion involves two ideas, space and time, — the distance moved over and the time taken. Consequently, the fundamental notions of every material phenomenon are mass, space, and time; and in describing any phenomenon or stating its law of action, certain definite amounts of each of these quantities are involved.

8. Fundamental Units. Before an amount of any quantity can be expressed mathematically, a unit, or standard, by which to measure it, must be selected. The following physical units have been adopted by the scientific world.

The centimetre. The unit of length is called the "centimetre," and it is the one hundredth portion of the length of a certain metal bar called the metre, which is kept in

Paris, the length of the bar being measured when it is surrounded by melting ice, i. e. at the temperature of 0° centigrade. The unit of area is the square centimetre; and the unit of volume is the cubic centimetre.

When this bar in Paris was constructed, it was designed to have a length equal to one ten-millionth of the distance from the equator to the pole of the earth along any meridian; but now that later measurements have shown that this relation between the bar and the earth is not exact, the scientific unit of length refers to the bar itself, and has no connection with the size or shape of the earth. Many copies of the standard metre-bar have been made; and although none are exact, their errors are accurately known.

The gram. The unit of mass, i. e. the unit amount of matter, is called the "gram;" and it is the one-thousandth portion of the mass of a certain piece of metal called the kilogram, which is kept in Paris. Many copies of the kilogram have been made; and although none of them are *exactly* correct, yet the differences between them and the standard have been most carefully measured.

When this standard mass in Paris was made, it was designed to be of such an amount that the mass of one-thousandth of it should be the same as that of one cubic centimetre of distilled water when it is densest, i. e. at the temperature of 4° centigrade. This relation has been shown to be not perfectly exact; and the standard of mass, then, is the gram itself, not the cubic centimetre of water. But for all ordinary practical purposes we may regard the mass of one cubic centimetre of water at 4° centigrade as being one gram.

The "density" of any homogeneous substance is defined as the number of grams in one cubic centimetre of the substance, and, in general, the density of any homogeneous portion of the substance is the mass of that portion divided by its volume. Thus the density of water at 4°

centigrade is one; and methods will be described later on for the measurement of the densities of all other bodies.

The second. The unit of time is called the "second;" and it is the second with reference to the so-called "mean solar day." That is, the second which is the scientific unit of time is the $\frac{1}{86400}$ portion of the average length of a solar day for one year.

The solar day is the interval of time which elapses between two successive transits of the sun across the meridian of the earth at the point of observation; and this interval is not the same at all times of the year. The average length of the intervals for one year is, however, constant from year to year, so far as we now know.

These units are, of course, perfectly arbitrary, and are not based on quantities which are in themselves constant. But they are most convenient, are simply defined and connected, and are in universal use. The system of measurement based upon them is called the "C. G. S. System," from the initial letters of the three standards; and the symbols ordinarily used are "cm.," "g.," "sec."

Methods for the measurement of lengths, masses, and intervals of time are fully described in laboratory manuals, and so need not be given here.

9. Mechanics. As has been said, all phenomena depending upon matter involve simply the idea of motion of matter. And so the study of Physics must begin with a discussion of the different possible kinds of motion, of the properties of matter in motion, and of the conditions under which these motions may occur. This discussion forms what is called the science of "Mechanics."

THEORY OF PHYSICS

BOOK I

MECHANICS AND PROPERTIES OF MATTER

CHAPTER I

KINEMATICS

KINEMATICS is that branch of Mechanics which is concerned with the different possible kinds of motion, and with a discussion of their distinguishing features.

10. Motion in General. If any body, such as a book or a stone, is tossed in the air at random, it is at once evident that there are two distinct motions: the body moves as a whole; and at the same time it is twisting. It is possible, of course, to have motion such that one only of these types occurs. Thus, if a book be raised from a table so that all lines which could be imagined drawn through the book remain parallel to themselves, the book simply moves as a whole, it does not twist. This motion is called pure "translation;" and it may be defined as such motion that all the points of the body move through equal parallel paths. Again, if some point of a body is fixed, the body cannot move as a whole, it can only spin or turn. An illustration of this is a spinning top whose point does not move, or a swinging door. In the case of the door, the points on the line through the hinges are at rest, while the other points of the door are all describing circles around this line. In the case of a spinning top, the geometrical axis may be considered at rest at any one instant, and all the points of the top are describing circles around it. If the axis of the top is inclined to the vertical, it does not *remain* at rest, but changes its direction. Such motion as

that of a door or a top is called "rotation;" and it may be defined as motion where at least one point of the body is at rest, while the others are moving at any instant in circular paths around a line, called the "axis," which passes through the fixed point. The axis may be fixed, as in an ordinary door or fly-wheel; or it may be changing its direction, as in a spinning-top when the axis of figure is not vertical. The most general type of motion is a combination of translation and rotation such as the motion of a nut on a screw.

It should be clearly understood that all motion is purely *relative.* If it is said that a stone falls in a straight line, it is meant that the path of the stone with reference to the earth is a straight line. The actual path of the stone in space will be entirely different from a straight line, owing to the motion of the earth itself, of the sun, etc. If in a railway-carriage running north, a passenger walks directly across the aisle from the east side to the west, his motion with reference to the carriage is west; with reference to the earth, it is northwest; with reference to the sun, it depends upon the hour and day; etc. It is impossible to conceive what the true motion is, because the human mind cannot imagine any point absolutely at rest. Similarly, in motion of rotation, the actual true motion cannot be conceived; and only the relative motion is studied. Ordinarily in Physics all motions are referred to the earth as at rest.

11. Translation. In translation all the points of a body move through equal and parallel paths; so that, if the change of position of any one point of the body is known, the motion of the whole body is also known. To describe and understand translation completely, then, the motion of a point must be studied; that is, its rate and direction of motion at successive instants must be known.

12. The change of position of a point is called the "displacement;" and the rate of this change with respect to

the time is called the "linear velocity." If this change is uniform, e. g. a railway train running at a constant rate, the velocity is the displacement in one second, and is the same for successive instants. If, however, the change of position is not uniform, e. g. a train slowing up, the velocity at any instant is the displacement which would occur in one second if the rate of motion remained for one second the same as it is at that instant. Thus, when a train is said to be going at the rate of sixty miles an hour, it is not implied that the train will actually go sixty miles in the following hour, but that it would do so if its rate were to remain unchanged.

If a point is displaced, two ideas are involved, — distance and direction. Thus, if the point moves from P to Q, the displacement is the line $\overline{PQ}$, which has a certain length, and the direction from P to Q. (This is often indicated by drawing an arrow-head on the line in the direction of the motion.) The line $\overline{PQ}$ is of course equal to $-\overline{QP}$, because if the point moves from P to Q, and then back to A, the displacement is zero. Hence $\overline{PQ}+\overline{QP}=0$. This idea of lines having direction as well as length should be familiar to all from a study of trigonometry. Since, then, displacement implies direction as well as distance, linear velocity must mean rate of motion in a particular direction. Thus, a train going north at the rate of 1000 centimetres per second is said to have the velocity "1000 north;" while a train going south at the same rate has a velocity "1000 south;" and the two velocities are equal and opposite, just as 5 and -5 are equal and opposite, or as the line $\overline{PQ}$ is equal and opposite to the line $\overline{QP}$. To express completely, then, the linear velocity at any instant, two quantities are necessary, — a number, to give the rate of motion, and a direction.

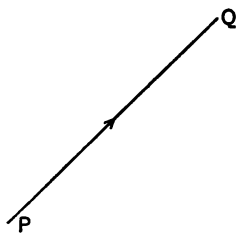


FIG. 1.

This number, which gives the rate of motion, without any statement of direction, is called the "linear speed." If the speed does not change, it represents the actual distance in centimetres which is passed over in one second; and the distance traversed in t seconds is the product of t and the speed. The speed at any instant is measured by determining how many centimetres are passed over in a certain time, and then dividing the distance by the time. Experimental methods are taught in laboratories.

Since a straight line has the same two properties as a velocity, — a number, giving its magnitude, and a definite direction, — it can be used to represent the linear velocity at any instant. Thus let the velocity at any instant be "10 north;" a straight line drawn from A to B in a northern direction, and having a length of ten arbitrary units, will completely represent the velocity. The length of the line gives the speed, and its direction must be the same as that of the displacement at that instant. Any other line, $\overline{A'B'}$, parallel to $\overline{AB}$ and of the same length, would do just as well, because being parallel they have the same direction, and they are made of the same length.

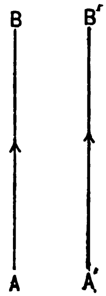


FIG. 2.

As an illustration of this mode of representation of linear velocities, consider the motion of a point in a circle, the speed being constant, e. g. a stone whirled in a sling, in a horizontal plane. When the point is at P, its direction of motion is along the tangent to the circle at P, i. e. perpendicular to the radius OP. Hence, if the speed is s, the velocity at P may be represented by a line, $\overline{AB}$, perpendicular to OP, and having a length equal to s arbitrary units. Similarly, at Q the velocity may be represented by a line, $\overline{A'B'}$, perpendicular to the radius OQ, and having a length s, because the speed remains constant.

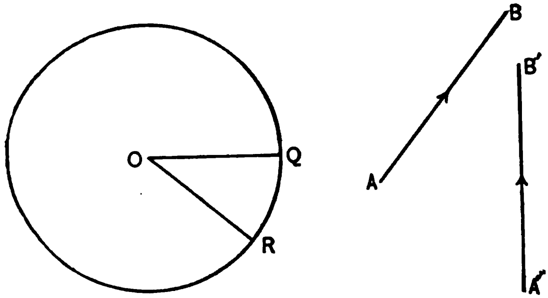


FIG. 3.

13. It often happens that a body's actual displacement is due to two or more causes, which by themselves individually would have produced different displacements. The actual displacement may, then, be regarded as the sum of the separate ones. Thus, when a man walks across a railway carriage which is in motion, his actual displacement with reference to the earth is made up of two parts, — his motion across the carriage, which may be represented by the line $\overline{AB}$, and the motion of the carriage with reference to the earth, which may be represented by the line $\overline{BC}$. So that, starting from the point A, the man reaches the point C with reference to the earth. If these distances are all passed over in one second, the lines $\overline{AB}$, $\overline{BC}$, and $\overline{AC}$ represent velocities; and it is thus proved that if a point is subjected to two velocities, $\overline{AB}$ and $\overline{BC}$, the actual velocity will be given, in direction and amount, by the line $\overline{AC}$. Similarly, consider a man rowing a boat across a river which has a strong current. Let $\overline{AB}$ be the velocity given the boat by the oars, and $\overline{BC}$ be that given it by the current; then the velocity of the boat at any instant will be represented by the line $\overline{AC}$, which completes the triangle,

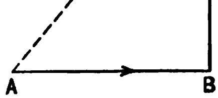


FIG. 4.

two of whose sides are $\overline{AB}$ and $\overline{BC}$. This process is sometimes spoken of as "composition of velocities," or "geometrical addition." In adding lines in this way, care must be taken to join two ends which are such that the arrows indicate continuous motion through the junction. The numerical value of $\overline{AC}$, or the actual speed, is found, by means of trigonometry, from a knowledge of the lengths of the other two sides and the angle between them. Using the symbol $\overline{AB}$ to mean the *line* A to B, *not* its mere length, the above facts may be written

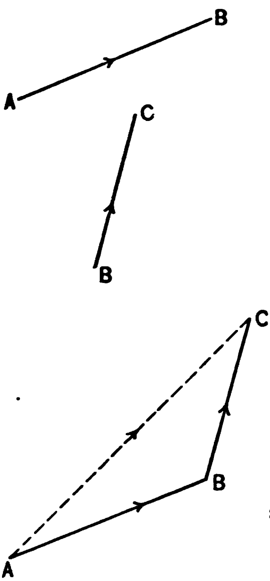


FIG. 5.

$$\overline{AC} = \overline{AB} + \overline{BC};$$

and from this it follows that

$$\overline{AC} - \overline{AB} = \overline{BC}.$$

The difference between two velocities is thus found by drawing the lines which represent them, both away from the same point, and then constructing the third side of the triangle.

If a point has more than two velocities given it simultaneously, the actual velocity may be found by adding two, then adding the third to their sum, and so on; or, what is the same thing, join all the lines which represent the velocities, in such a way that the arrows point in a continuous direction, and then draw a line to complete the polygon. This will be the actual velocity.

14. If the motion of translation is uniform, it is completely defined by a knowledge of the linear velocity; but in all actual cases the velocity is changing, and so a knowledge of the nature and amount of this change is required. The name "linear acceleration" is given to the

rate of change of the linear velocity with reference to the time; and it may be uniform or not. If the change is uniform, the acceleration is the change in the velocity in one second; if it is not uniform, the acceleration at any instant is the change in the velocity which would take place in one second if the rate of change did not vary in that time. Since velocity has two essential properties, direction and amount (i. e. speed), it may be changed in two ways, either in direction or in speed. Thus there are two types of linear acceleration: 1. Where the direction is unchanged, but the speed is altered, e. g. a falling body; 2. Where the speed is unchanged, but the direction is altered, e. g. a stone whirled in a sling in a horizontal plane. In the general case, both speed and direction are changing.

An acceleration can also be represented by a straight line, because, being the rate of change of velocity, it is the difference between two velocities. It has both a definite direction and a definite amount.

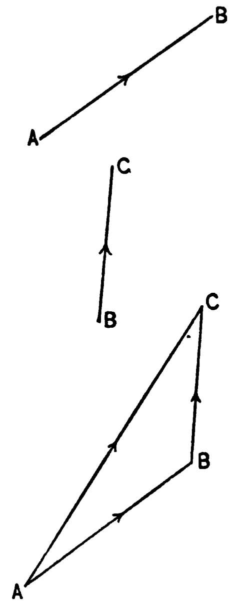


Fig. 6.

15. Accelerations can be added, then, or compounded. Thus, let a point be subject to two accelerations, $\overline{AB}$ and $\overline{BC}$; the actual acceleration will then be the line $\overline{AC}$, which is the geometrical sum of $\overline{AB}$ and $\overline{BC}$. Looking at this theorem in the reverse way, it may be said that any actual acceleration may be regarded as the sum of any two accelerations which when added geometrically equal it. There is thus an indefinite number of ways in which an acceleration may be broken up, or "resolved," into parts. In the figure the acceleration $\overline{AC}$ is regarded as equiva-

lent to an acceleration $\overline{AB}$ in the direction A to B, and $\overline{BC}$ in the direction B to C. But this last acceleration also produces a certain effect in the direction A to B. So that, if $\overline{AC}$ is to be resolved into two mutually independent accelerations $\overline{AB}$ and $\overline{BC}$, these last must be at right angles to each other. In this case the *total* effect in the direction A to B of the acceleration $\overline{AC}$ is given by the line $\overline{AB}$; and the total effect in the direction B to C is given by the line $\overline{BC}$. For this reason the acceleration $\overline{AB}$ is called the resolved portion of the acceleration $\overline{AC}$ in the direction A to B. If θ is the angle between the lines A to C and A to B,

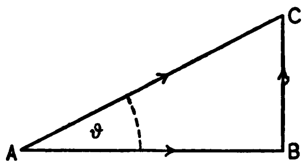


FIG. 7.

$$\overline{AB} = \overline{AC} \cos \theta,$$

$$\overline{BC} = \overline{AC} \sin \theta.$$

An illustration of this is afforded by a body sliding down an inclined plane. Let the angle between the plane and the horizontal plane be ϕ, and let the body be at A at any instant. If the plane were not there, the body would experience an acceleration $\overline{AC}$ vertically downwards. $\overline{AC}$ can be regarded as producing the accelerations $\overline{AB}$ and $\overline{BC}$ perpendicular and parallel to the plane respectively; but now, if the body rests on the plane, the acceleration $\overline{AB}$ perpendicular to the plane is prevented, and the actual acceleration will be $\overline{BC}$ parallel to the plane. Further $\overline{BC} = \overline{AC} \sin (BAC) = \overline{AC} \sin \phi$. This deduction

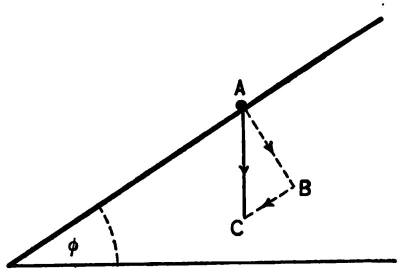


FIG. 8.

may be verified by experiment, as both $\overline{AC}$ and $\overline{BC}$ can be measured.

In an exactly similar way a velocity may be resolved into components in directions at right angles to each other.

16. Rotation. In rotation, at least one point of a body is at rest, and all the other points are at any instant moving in circles whose centres all lie on a line which passes through the fixed point. Consider any plane section of the body perpendicular to the axis; draw lines AA' and BB'; in this plane of the body, also draw a line, PP', lying in this plane, but fixed in space. Let O be the point where the plane cuts the axis; then, if the axis is fixed, the lines AA' and BB' will make certain angles with the fixed line PP', which will change with time as the body rotates. But the change in the angle made between AA' and PP' equals the change in the angle made between BB' and PP' because AA' and BB' have a constant angle between them. This is equivalent to saying that all lines, AA', BB', etc., fixed in the body turn through equal angles with reference to any line, PP', fixed in space, these lines all being drawn in a plane perpendicular to the axis of rotation. Consequently, just as in motion of translation all points of the body move through the same distances in the same direction, so in motion of rotation all lines in a plane perpendicular to the direction of the axis have the same angular motion around that axis. The angle between the line AA' fixed in the body and PP' fixed in space determines the position of the body; and the rate of change of this angle with respect to the time is called the "angular velocity." It is evident

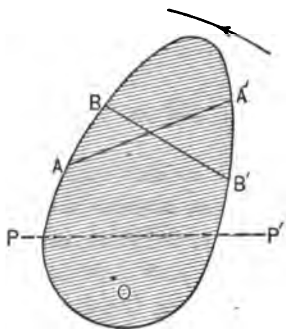


FIG. 9.

that two quantities are necessary for its complete definition; the direction of the axis so that a plane may be drawn perpendicular to it, and the numerical value of the rate of angular motion, i. e. the "angular speed." Consequently an angular velocity may be represented by a straight line. Thus $\overline{AB}$ may represent the turning of a body whose axis at the instant has the direction A to B, and whose rate of turning at the instant equals the numerical length of $\overline{AB}$. A line $\overline{BA}$ would represent a body turning with the same speed around the same axis, only in the opposite direction. And just as linear volocities may be added, so may angular ones, if the axes of rotation of the different motions meet in a point.

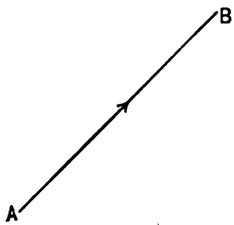


FIG. 10.

In translation either the direction or the speed may change, and thus produce linear acceleration; so, in rotation, either the direction of the axis may change, as in the case of a spinning top whose axis of figure is not vertical, or the angular speed may change, the direction of the axis remaining unaltered, as in the case of a barrel rolling down an inclined plane. Thus there are two types of angular acceleration, and the general case is a combination of these two; that is, both the direction of the axis and the angular speed around the axis change. Since angular acceleration is the rate of change of angular velocity, it can be represented by a straight line; and angular accelerations can be added or resolved into components around particular axes, just as linear accelerations can be added or resolved into components in particular directions.

17. Numerical Value of an Angle. It may be well, at this point, to recall the scientific definition of an angle and its numerical value. An angle between two lines lying in the same plane is the difference in direction between them; and its numerical value is defined in this way: prolong

the lines until they intersect, then with any radius, r, describe a circular arc around this point of intersection as a centre, having the arc terminated by the two lines. Let the length of the arc be a; then the numerical value of the angle is a / r. This value is independent of the length of the radius which is taken, as appears from ordinary geometry.

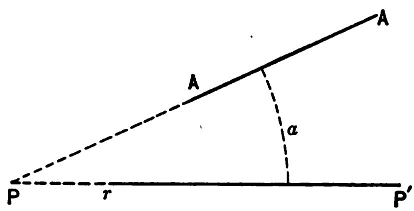


FIG. 11.

As illustrations: a right angle has the numerical value $\frac{2 \pi r}{4} / r = \frac{\pi}{2}$; one entire turn corresponds to an entire circumference, and has the numerical value $2 \pi r / r = 2 \pi$.

SPECIAL CASES

I. TRANSLATION

18. 1. Motion in a Straight Line; Constant Acceleration. The problem is to determine the velocity and the distance passed over at the end of a certain time.

Let s_0 be the initial speed, that is, the speed at the instant from which time is counted; a, the acceleration, which in this case is a known constant quantity; t, the number of seconds at the end of which it is desired to know the velocity and the distance traversed. Since in this problem the motion is in a straight line, the acceleration is the numerical change in the speed. Hence at the end of t seconds the speed will be

$$s = s_0 + a t \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Since the speed is changing uniformly, the mean speed during the t seconds is $\frac{s_0 + s}{2}$, the numerical average of

the initial and final values. This is the average distance passed over in one second; and so the entire distance traversed in t seconds is

$$x = t \cdot \frac{s_0 + s}{2}.$$

Substituting for s its value from (1),

$$x = s_0 t + \tfrac{1}{2} a t^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Eliminating t from (1) and (2),

$$s^2 - s_0^2 = 2\,a\,x \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Conversely, if by observation it is found that the motion is such that $x = \dfrac{s^2 - s_0^2}{2\,c}$, where c has a constant numerical value for all values of x, then it is known that the acceleration is constant, and c is its numerical value. This is a method of measuring a constant acceleration.

As illustrations of these formulæ, consider two examples.

a. A body falls from rest towards the earth. Experiments prove that falling bodies have a constant acceleration towards the earth of about 980. Hence, since $s_0 = 0$ and $a = 980$,

$$s = 980\,t;$$
$$x = \tfrac{1}{2} \cdot 980 \cdot t^2$$
$$s^2 = 2 \cdot 980\,x.$$

If $\quad t = 1, \; s = 980, \quad x = 490,$

if $\quad t = 2, \; s = 1960, \; x = 1960.$

b. A body is thrown vertically upwards with an initial velocity 4900. How far will it go before it comes to rest? $s_0 = 4900$; $a = -980$, since in this case the acceleration is opposite to the speed; also $s = 0$, since the final velocity is to be 0.

Hence
$$0 = 4900 - 980\,t;$$
$$x = 4900\,t - 490\,t^2;$$
$$0 - 4900^2 = -2 \cdot 980 \cdot x.$$

Hence $\quad t = 5; \; x = 12250.$

19. 2. Motion in a Circle; Constant Speed.

The problem is, knowing the speed, to determine the acceleration, — both its direction and its amount. Let the motion be in a circle of radius, r, around a centre, O; and let P and Q be two points of the circle, which are so near that they may be regarded as consecutive; and let it take the extremely small interval of time t seconds for the point to move from P to Q (t may be 1 millionth, or even smaller). Then, if s is the speed, $\overline{PQ} = s \cdot t$ because s is the distance passed over in one second.

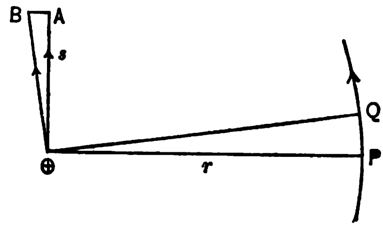


FIG. 12.

The *velocity* at the point P may be represented by the line $\overline{OA}$, whose length is s, and which is perpendicular to $\overline{OP}$. Similarly, the velocity at Q is given by the line $\overline{OB}$, whose length is s, and which is perpendicular to $\overline{OQ}$. Hence the change in the velocity in t seconds is $\overline{OB} - \overline{OA}$, that is, the line $\overline{AB}$. But a, the acceleration, is the change in the velocity in one second; hence

$$\overline{AB} = a \cdot t.$$

Hence, eliminating t from the two equations,

$$\frac{a}{s} = \frac{\overline{AB}}{\overline{PQ}}.$$

But since the angles AOB and aOb are equal,

$$\frac{\overline{AB}}{\overline{PQ}} = \frac{\overline{OA}}{\overline{OP}} = \frac{s}{r}.$$

Hence $$\frac{a}{s} = \frac{s}{r};$$

and so $$a = s^2 / r \quad \ldots\ldots \quad (4)$$

This is the numerical value of the acceleration when the point is at P; its direction is given by the line $\overline{AB}$ when the time t becomes infinitely small. In the limit, $\overline{OQ}$ coincides with $\overline{OP}$; and the line $\overline{AB}$ is *parallel* to them. That is, the acceleration at the point P has for its numerical value s^2/r, and is directed toward the centre of the circle. If this acceleration ceased being applied, the point would move off in a tangent to the circle at that instant. This explains the breaking of fly-wheels when the speed becomes too great, and also the tendency of all matter to place itself as far as possible from the centre when it is whirled around an axis.

As the point moves around the circle, the radius joining it to the centre turns through a constantly increasing angle with reference to any fixed line. The angular speed of this line is the angle turned through in one second, but an angle has for its numerical value the ratio of the arc to the radius. So that, choosing the radius as that of the circle r, the arc passed over in one second is the linear speed s. Hence, calling the angular speed ω,

$$\omega = s \,/\, r \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (5)$$

And substituting for s its value in terms of ω, (4) becomes

$$a = r\,\omega^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (6)$$

Call n the number of complete turns which the point makes around the circle in one second.

$$s = n \cdot 2\,\pi\,r \quad \text{and} \quad \omega = n \cdot 2\,\pi.$$

Hence $$a = 4\,\pi^2\,n^2\,r \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (7)$$

The period of time for one complete revolution around the circle is $\frac{1}{n}$; that is, $2\,\pi\,r \,/\, s$ or $2\,\pi \,/\, \omega$; and calling its value T, the formula for the acceleration becomes, on its substitution,

$$a = 4\,\pi^2\,r \,/\, T^2 \quad . \quad . \quad . \quad . \quad . \quad . \qquad (8)$$

20. Comparing these two special cases, 1 and 2, it is seen that the speed is changed by applying an acceleration in the direction of motion; while the direction is changed by applying an acceleration at right angles to the direction of motion. If an acceleration is applied obliquely to the direction of motion, it may be considered replaced by its two components parallel and perpendicular to this direction; and consequently both the speed and direction will be changed.

21. 3. **Harmonic Motion.** This is a motion which may be thus described: Let a point, P, move around a circle of radius, r, with a constant speed, and take at consecutive instants the projection of this point on any diameter of the circle. Then, as this projected point, Q, moves backward and forward along the diameter, it has a certain kind of vibratory motion, which is called "harmonic." The great importance of this kind of motion is due to the fact that it occurs so often in nature, and that all instruments for measuring time accurately have motions like it.

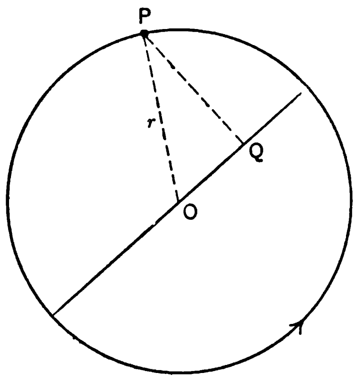


FIG. 13.

The point Q has a certain acceleration along the diameter, and the problem is to determine the connection between this acceleration and the period of one complete vibration of Q.

The acceleration of Q towards O along the diameter is the same as the component of the acceleration of P parallel to the diameter, because Q is the projection of P on it. The acceleration of P towards O is by (6) $r\,\omega^2$ if ω is the constant angular speed of P. Hence the component of

this acceleration parallel to the diameter is $r\,\omega^2 \cdot \cos(P\,O\,Q)$, and this is also the acceleration of Q towards O along the diameter. Call the line $\overline{O\,Q}$ x, where x, of course, is a variable quantity, its greatest value being $+r$ and its least $-r$. Then $\cos(P\,O\,Q) = x/r$, and the acceleration of Q towards O may be written

$$r\,\omega^2 x\,/r, \quad \text{or} \quad x\,\omega^2;$$

i. e., at any point the acceleration equals a constant times x.

The period of a complete vibration of Q is the same, obviously, as the period of revolution of P around the circle. Hence

$$T = 2\,\pi\,/\,\omega = 2\,\pi\,/\,\sqrt{\text{coefficient of } x \text{ in the acceleration.}} \quad (9)$$

Therefore, simple harmonic motion is such motion that at any instant the acceleration is towards the centre of the path, and has for its value a constant times the distance of the point from the centre; and, further, the period of vibration is equal to $2\,\pi$ divided by the square root of this constant. The length of the path (i. e. the diameter of the circle) is called the "amplitude" of the vibration. Illustrations of harmonic motion are afforded by the vibrations of pendulums, tuning-forks, stretched strings, etc.

II. Rotation

22. 1. Motion around an Axis whose Direction does not change; Angular Acceleration Constant.

This is perfectly analogous to Case 1, in Translation. Let ω_0 be the initial angular speed, and a the angular acceleration; then, at the end of t seconds, the speed ω, and the angle turned through, θ, will be

$$\omega = \omega_0 + a\,t \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

$$\theta = \frac{\omega_0 + \omega}{2} \cdot t = \omega_0\,t + \tfrac{1}{2}\,a\,t^2{}_0 \quad . \quad . \quad (11)$$

and $$\omega^2 - \omega_0^2 = 2\,a\,\theta \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

As illustrations of this motion, consider a cylinder rolling down an inclined plane and a fly-wheel being stopped by a brake applied at its rim.

23. 2. **Motion of a Spinning Top whose Axis of Figure is not Vertical.** The axis changes its direction at a uniform rate, and the angular speed around the axis is constant.

This is perfectly analogous to Case 2 in Translation. In that problem it was shown that in order to make the point change its direction at a uniform rate (i. e. to make it move in a circle) the speed remaining constant, it is necessary to apply continuously to the point a constant acceleration whose direction is perpendicular to the direction of motion. So in this problem of rotation, in order to make the axis change its direction at a uniform rate, the angular speed remaining constant, it is necessary to apply a constant angular acceleration around an axis which is perpendicular to the axis of rotation. In the case of a spinning top there is a constant tendency to fall over, i. e. to turn around a horizontal axis; and this axis is perpendicular to the axis of rotation, hence this last axis changes its direction at a uniform rate.

24. In general, then, if a body is spinning with the angular velocity $\overline{AB}$, the angular speed can be changed by applying an angular acceleration around the same axis; but if the direction of the axis is to be changed, an angular acceleration, $\overline{AC}$, around an axis perpendicular to the original axis and intersecting it must be applied. The resulting angular velocity will be the sum of $\overline{AB}$ and $\overline{AC}$ or $\overline{AD}$; and the new axis obviously lies in the same plane as the original axis and the axis of the acceleration. Unless there is, then, an *angular* acceleration, neither the angular speed nor the direction of the axis of rotation

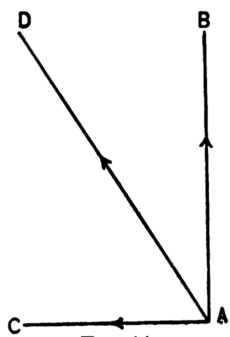


FIG. 14.

changes. So that if it is desired to have a body move in such a way as to keep some plane in it always in the same direction, all that is necessary is to give it a rotation around an axis perpendicular to this plane. For, then, unless an *angular* acceleration is impressed, this axis will keep its direction fixed. Thus a quoit thrown at random will turn over and over; but if it be given a twist as it is thrown, it will move with the axis remaining parallel to itself. There are many other illustrations, such as the twist given projectiles from guns, the twist given knives when tossed by a juggler, etc.

25. 3. **Harmonic Motion.** There is also harmonic motion of rotation, as is illustrated by the balance-wheel of a watch, a magnetic compass-needle when slightly displaced, a wire one of whose ends is clamped and whose other end is twisted and then allowed to vibrate. For rotation, it may be defined as such motion that the angular acceleration at any instant is always towards the central position, and its numerical value is a constant times the angle through which the body is turned; and, further, the period of vibration equals 2π divided by the square root of this constant. The "amplitude" is the angle which marks the extent of the vibration.

There are, of course, many other special kinds of motion; but these are perhaps the most important.

CHAPTER II

DYNAMICS

DYNAMICS is that branch of Mechanics which is concerned with the properties of matter in motion and with the conditions under which this motion may be produced or changed.

26. Measurement of Mass by means of Inertia. That property of matter which is most familiar to us, perhaps, is its inertia. As already explained, this is a name given to that property which becomes evident to our muscular sense when we change the motion of matter in any way. If we push a barrel along a floor, or if we put a ball in motion, we are conscious of a definite sensation, which is said to be due to the inertia of the matter. We can estimate in a rough way the amounts of matter in different bodies by the sensations produced owing to their inertia; and we can easily tell whether a box is empty or full by trying to push it. We will, then, define two bodies as having the same mass if they have the same inertia. (We might perfectly well define equal masses as corresponding to equal "weights;" and we have no right, *a priori*, to regard these two definitions as being identical, although they may be proved to be so by experiment.) So, if our senses were delicate enough, we might perfectly identify two equal masses. But since they are not so delicate, some instrument may be devised which is much more sensitive and which can produce the same effect as our muscles. Such an instrument is a spiral spring when compressed; it can produce change of motion, and can

be tested most carefully. By means of it, it is theoretically possible to determine whether two bodies have the same inertia, that is, the same mass on the above definition of equal masses. The experiment might be performed thus: compress the spiral spring to a certain definite point; place one of the bodies on one end of the spring, keeping the other end firmly clamped; allow the spring to expand, and accurately measure the velocity given the body after it leaves the spring. (It is necessary to have the spring so arranged that the body is projected along a perfectly smooth horizontal table, so as to do away with all effects of friction and "weight.") Then compress the spring to identically the same point as before, place the second body upon it, and measure the velocity given it when it is projected. In general, these two velocities would be different; but it is theoretically possible to so change one of the bodies by chipping off small pieces that the velocities given the two are the same, within the range of the inherent errors of the experiment. These two bodies would then have equal masses as closely as they could be measured. Combining these two, it would be possible to make a third body have the same mass as their combined mass, that is, twice the original mass; and so the process might be continued until a series of bodies was obtained whose masses were 1, 2, 3, 4, 5, etc. times that of the original body. By taking the original mass as small as is wished, it is perfectly possible thus to make standards of mass varying by small steps from very small masses to very large; and the mass of any other body can be determined by finding what combination of the standard masses has the same inertia as the body, using a delicate spiral spring for the comparison. The scientific standard of mass is the gram, as already defined; and as it can be compared with this arbitrary set of masses, their values in terms of the gram may be easily deduced; and so all measurements of masses may be given in grams.

Matter in Translation

27. Since in translation all the points of the body have the same motion, the entire matter may be considered as concentrated at one point; and this point thus endowed with inertia may be called a "particle."

There is one general law in regard to the translation of matter, which so far as tested by experiment seems to be absolutely true and to have no exception. It is this: consider a number of bodies of all kinds in motion of translation; and let their motion be *entirely free from all external influences;* they will impinge on each other or will affect each other in some way, so that their velocities will be constantly changing; let their masses be m_1, m_2, m_3, etc., and let their linear velocities at any instant be u_1, u_2, u_3, etc.; form the *geometrical* sum $m_1 u_1 + m_2 u_2 + m_3 u_3 +$ etc.; then the law is that this sum is the same for all other instants, no matter how the individual velocities change. That is, the geometrical sum

$$m_1 u_1 + m_2 u_2 + m_3 u_3 + \text{etc.} = \text{a constant} \quad . \quad (1)$$

The geometrical sum must be taken, because the velocities may be in different directions, and so cannot be added algebraically. The law may also be stated in this way: let w_1, w_2, w_3, etc. be the components of the linear velocities in any fixed direction; then the algebraic sum

$$m_1 w_1 + m_2 w_2 + m_3 w_3 + \text{etc.}$$

is the same at all times, if there are no external influences.

This law cannot be verified directly in the most general case, but certain special cases of it may be; and the law is regarded as true in general, because every deduction from it that admits of experimental verification has been found to be true.

Owing to the importance of the product $m u$, mass times linear velocity, it has been given a name, "linear momen-

tum;" and this general law of nature is called the "Principle of the Conservation of Linear Momentum."

Special Cases of Translation. **28.** 1. Let there be only one body. The law states that, if there are no external influences,

$$m\,u = \text{a constant} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

But the mass of a body cannot change; that is, m is a constant. Hence the velocity of translation must remain constant. It follows, then, that a body which has at any instant a certain speed in a certain direction will continue to move with the same speed in the same direction, if it is free from all external influence. This seems to be perfectly in accord with experiments. A special case of this is when the body is observed to be at rest at any instant. Since the velocity is zero at that instant, it must be zero for all other instants; or a body once at rest will always remain at rest, if left to itself.

29. 2. Let there be two bodies. Then, if there is no external action, that is, if these bodies are influenced only by each other,

$$m_1 u_1 + m_2 u_2 = \text{a constant} \quad . \quad . \quad . \quad . \quad (3)$$

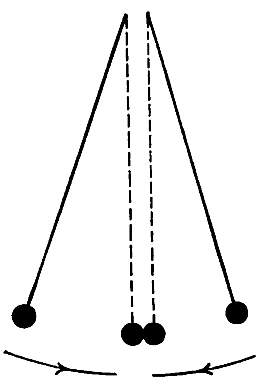


Fig. 15.

A special case of this deduction may be verified in this way: —

Suspend two spheres by cords so that they may each swing freely in vertical circles. Make the radii of these circles the same length, and hang the spheres side by side. Now, if they are drawn one side and then allowed to impinge, each will have a certain velocity before impact and a definite velocity immediately after. These velocities are in the same straight line, if the bodies are spheres (and also in other cases if the bodies are of suit-

able shapes); and, as will be shown later on, they can be easily measured. The masses can also be measured; and so it can be proved that the sum $m_1 u_1 + m_2 u_2$ is the same before and after impact, although the individual velocities change.

Another experiment which also verifies the law is this: if a man standing on a board which rests on a smooth table jumps off, the board goes in the opposite direction; and if m_1 is the mass of the man, and u_1 the velocity with which he jumps, while m_2 and u_2 are the mass and velocity of the board, then it is found that

$$m_1 u_1 = - m_2 u_2.$$

This is in agreement with the law; because, before the man jumps, he and the board are at rest, and hence the two velocities are at that instant zero. So the sum

$$m_1 u_1 + m_2 u_2$$

must always be zero, if the only action is one between the board and the man. Consequently,

$$m_1 u_1 + m_2 u_2 = 0,$$

or

$$m_1 u_1 = - m_2 u_2,$$

which means that the velocity of the board is opposite to that of the man, and that its numerical value is $\frac{m_1}{m_2} u_1$.

An exactly similar illustration is afforded when a bullet is fired from a gun, or when a body falls towards the earth.

The illustrations of this law might be multiplied indefinitely; but it is sufficient to say that no exception to it is known, and hence it is believed to be one of the great laws of nature.

30. In terms of momentum, this law may be thus stated: the sum of the linear momenta of any number of bodies in *any* fixed direction is a constant, provided that no influence external to the system is acting on it.

In general, however, there is some external action; and then the momentum may change. Thus, owing to the presence of the earth, there is always a change in the momentum of any piece of matter which is moving near it; a piece of iron near a magnet has its momentum changed, if it is free to move, owing to the presence of the magnet; a feather moving through the air has its momentum changed, owing to the presence of the air. The rate of change of the linear momentum of any body with respect to the time is called the "force acting on it." Linear momentum is $m u$; but the mass cannot change, and the rate of change of the linear velocity is the linear acceleration; consequently, force is the product of mass and linear acceleration. Therefore the mathematical expression for force is

$$f = m a \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

A unit force is called a "dyne." Thus, if $m = 1$, and $a = 1$, $f = 1$; that is, if the mass is 1 gram, and if in 1 second the velocity is changed by an amount 1 centimetre per second, the force is one dyne. As an illustration, it is found by experiment that, when a body falls towards the earth, the acceleration is nearly 980, i. e. that in each second the velocity is increased by nearly 980 centimetres per second; hence, if a mass 10 grams falls with this acceleration, the force is 9800 dynes.

If the rate of change of momentum is uniform, the force is the change of linear momentum in one second. Sometimes, however, it is impossible to measure the *rate* of change of momentum, e. g. when a body is struck a sudden blow; but it is always possible to measure the *total* change; and this total change of the linear momentum is called the "linear impulse."

Both forces and impulses have definite directions and definite numerical values; and so they can be represented by straight lines. Thus forces can be added, subtracted, and resolved into components. It may happen that a body

has two forces, as a piece of iron near the earth and near a magnet. Let one force be represented by $\overline{AB}$, the other by $\overline{AC}$; the actual force will be their sum, $\overline{AD}$. And if $\overline{AB}$ and $\overline{AC}$ are at right angles to each other, $\overline{AB}$ is the component of $\overline{AD}$ in the direction of A to B.

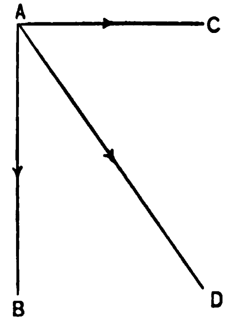


FIG. 16.

31. There is an acceleration produced in the motion of a body when it is allowed to fall towards the earth, that is, there is a force; and so sometimes it is said that the earth "exerts a force on" the body, or that the body is acted on by the force of the earth. Similarly, a magnet is said to exert a force on a piece of iron. So we speak of the "force of the earth's gravity," the "magnetic force," etc. But these expressions simply mean that under definite conditions the linear momentum of a body is changed by a definite amount per second. When a body is "acted upon by" two forces, each force produces its own effect; and the actual force is the geometrical sum of the two, as proved above. Similarly for three or more forces.

Another illustration is afforded by a body which is suspended from a vertical spiral spring. The spring is stretched by the body; but, since everything is at rest, we say that the force of the earth down is balanced by the force of the spring up; meaning that if the spring were not there, there would be a definite force, $m\,a$, downward; and if the earth were removed, there would be a force upward, $-m\,a$; and so the total force is zero. A spiral spring can thus be used to compare and measure any forces. Similarly, if a body is at rest on a table, the force down is balanced by the resistance of the table.

32. The laws of motion for one or two bodies, as above explained, may now be expressed in terms of forces.

1. The linear momentum of a body will not change, if there is no force "acting on" it.

2. If there is a force, it is measured by the rate of change of linear momentum. Further, each force produces its effect independently of the others.

3. If there are two bodies, whose linear momenta are changed owing to the presence of each other, the change in the momentum of the one must be equal but opposite to the change in the momentum of the other, because the sum of the two momenta does not change, and so the total change must be zero. Hence it may be said that the force which the first exerts on the second is equal and opposite to the force which the second exerts on the first; or, in other words, "action and reaction are equal and opposite." These statements are called "Newton's Laws of Motion."

Since force is mass times linear acceleration, there are two types of forces corresponding to the two types of linear acceleration: in one the speed is changed, not the direction; in the other the direction is changed, not the speed.

Special Cases. **33.** 1. Any body near the earth is "acted upon by" a force, mg, where m is the mass and g is a quantity which is constant for any one place and nearly equals 980.

Consider the motion of two bodies whose masses are m_1 and m_2, and which are connected by a cord; the first of these hangs vertically, the second rests on a smooth table, the cord passing over a pulley, as shown. There are two forces "acting on" the body whose mass is m_1; one is due to the earth, the other to the cord. Call the force due

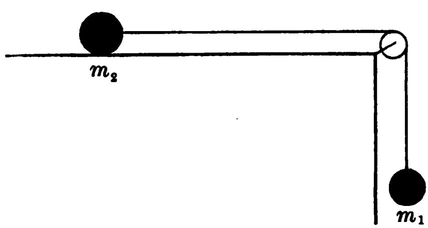


FIG. 17.

to the cord T; it evidently opposes the force $m_1 g$ due to the earth. Hence the actual force is $m_1 g - T$. That is,

$$m_1 a = m_1 g - T, \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where a is the actual acceleration.

The forces "acting on" the body whose mass is m_2 are the force due to the earth, the resistance of the smooth table to this force, and the tension or pull of the cord. The first two neutralize each other; the third equals T, because "action and reaction are equal and opposite." Further, m_2 must have the same acceleration as m_1, since they are connected by the cord. Hence

$$m_2 a = T \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Combining (5) and (6), it follows that

$$a = \frac{m_1}{m_1 + m_2} g \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

That is, the acceleration is a constant, and is in the direction of motion. Hence Formulæ (1), (2), (3) of Chapter I. can be applied to find the resulting speed and distance travelled.

34. 2. Let two bodies whose masses are m_1 and m_2 be connected by a cord which passes over a pulley; and let the bodies hang vertically. If m_1 is greater than m_2, the first body will move down; the second up.

The total force "acting down on" m_1 is $m_1 g - T$. Hence

$$m_1 a = m_1 g - T \quad . \quad . \quad . \quad (8)$$

The total force acting up on m_2 is $T - m_2 g$.

Hence

$$m_2 a = T - m_2 g \quad . \quad . \quad . \quad (9)$$

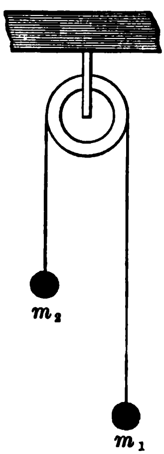


FIG. 18.

since the acceleration of m_2 upward equals that of m_1 downward. Combining (8) and (9),

$$a = \frac{m_1 - m_2}{m_1 + m_2} g \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

And the resulting speed and distance travelled may be at once calculated.

35. 3. Let two bodies, whose masses are m_1 and m_2, be connected by a cord and be made to revolve around each other by means of a "whirling-table." This is a piece of apparatus in which the two bodies are carried by a horizontal metal rod which can be made by suitable wheels and gearing to revolve rapidly around a vertical axis. The two bodies have holes through their centres, through which the rod passes; and so, when the rod revolves, the bodies turn around each other, the radii of their paths depending upon where the bodies are placed on the rod. In general the bodies will not continue their revolution, but will slip to one end or the other of the apparatus. It is always possible, however, to find some position in which they will not slip. Let the axis of revolution be at the point O; and call the distances from O to the centres of the two bodies r_1 and r_2, respectively.

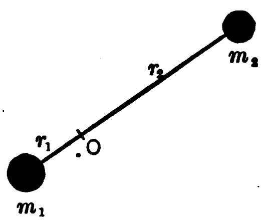


FIG. 19.

Since the body whose mass is m_1 is moving in a circle of radius r_1 with a certain angular speed, ω, its acceleration is $r_1 \omega^2$ towards the centre O. Hence the force is $m_1 r_1 \omega^2$. Similarly, the force "acting on" m_2 is $m_2 r_2 \omega^2$, since both bodies are moving with the same angular speed. But action equals reaction;

hence $$m_1 r_1 \omega^2 = m_2 r_2 \omega^2,$$

or $$m_1 r_1 = m_2 r_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

This may be easily verified by experiment.

36. 4. As explained above, since force is mass times acceleration, forces may be added and may be also resolved into components. Thus, consider two bodies whose masses are m_1 and m_2 connected by a cord which passes over two fixed pulleys. In between the pulleys let a body whose mass is m_3 be so suspended that it is free to slide along the cord. There is a force, $m_1 g$, "acting on" m_3 in the direction of the cord which passes over the first pulley,

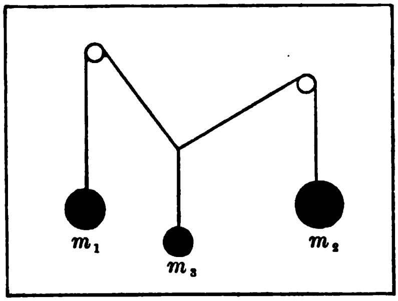


FIG. 20.

another force, $m_2 g$, in the direction of the second pulley, and also a force, $m_3 g$, vertically downward. Let these three forces be represented by the lines $\overline{OA_1}$, $\overline{OA_2}$, $\overline{OA_3}$. $\overline{OA_1}$ and $\overline{OA_2}$ combine to produce the force $\overline{OB}$. If this is equal and opposite to $\overline{OA_3}$, the three forces will neutralize one another; and there will be no acceleration. Consequently, if the body whose mass is m_3 is allowed to come to rest on the cord, the directions of the two sections of the cord must be such that the geometrical sum of $\overline{OA_1}$ and $\overline{OA_2}$ is exactly equal and opposite to $\overline{OA_3}$. This may be verified by experiment.

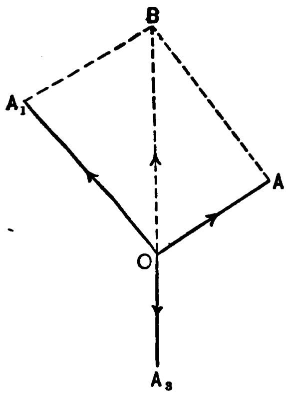


FIG. 21.

Another way of regarding this problem is as follows: Let the directions of the forces be as shown, when the mass m_3 comes to rest. The force $\overline{OA_1}$ can be resolved into two components, one in the direction of the line $\overline{OA_3}$

reversed, and the other at right angles to it; i. e. the two components are $\overline{OP_1}$ and $\overline{P_1A_1}$. Similarly the force $\overline{OA_3}$ can be resolved into two components, $\overline{OP_2}$ and $\overline{P_2A_2}$. Then, since the point O does not move,

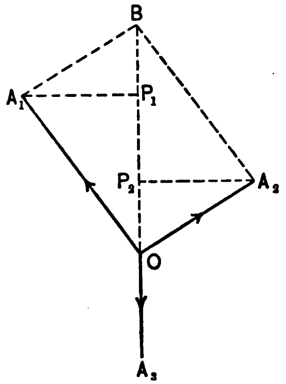


FIG. 22.

$$\overline{P_1A_1} = -\overline{P_2A_2}$$
$$\overline{OA_3} = -(\overline{OP_1} + \overline{OP_2}).$$

But these conditions are satisfied only if the parallelogram formed on the two lines $\overline{OA_1}$ and $\overline{OA_2}$ is such that the diagonal from O is equal and opposite to $\overline{OA_3}$.

37. 5. Consider the motion of two bodies whose masses are m_1 and m_2, and which are connected by a cord passing over a pulley, so that one body hangs vertically and the other rests on a smooth plane inclined to the horizon at an angle θ. The force of the earth on m_2 acts vertically down; its component parallel to the plane is $m_2 g \sin\theta$. Hence, if T is the tension of the string, the actual force on m_2 down the plane is

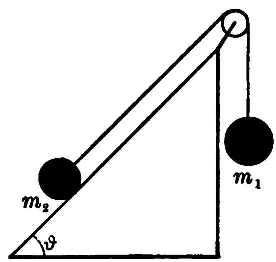


FIG. 23.

$$m_2 a = m_2 g \sin\theta - T,$$

and the force on m_1 is

$$m_1 a = T - m_1 g.$$

Hence
$$a = \frac{m_2 \sin\theta - m_1}{m_1 + m_2} g \quad . \quad . \quad . \quad . \quad . \quad (12)$$

38. Centre of Inertia. Let the masses of a system of particles be m_1, m_2, m_3, etc.; and let their distances at any instant from a fixed plane be x_1, x_2, x_3, etc. A special case

of this is, of course, a single body like a stone or a drop of water, where m_1, m_2, etc., are the masses of the small portions into which the body may be considered divided. Form the algebraic sum

$$m_1 x_1 + m_2 x_2 + m_3 x_3 + \text{etc.}$$

and divide this by $m_1 + m_2 + m_3 + \text{etc.}$, i. e. by the mass of the system. This quotient gives obviously the average distance of the entire system from the fixed plane. Call this distance $\bar{x}$. Then

$$(m_1 + m_2 + m_3 + \text{etc.})\,\bar{x} = m_1x_1 + m_2x_2 + m_3x_3 + \text{etc.} \quad (13)$$

Imagine these masses all in motion away from the fixed plane; each will have a definite velocity of its own which is the rate of change of its distance from the plane. Writing $\bar{u}$ for the rate of change of $\bar{x}$, u_1 for the rate of change of x_1, etc. (13), gives at once

$$(m_1 + m_2 + m_3 + \text{etc.})\,\bar{u} = m_1u_1 + m_2u_2 + m_3u_3 + \text{etc.} \quad (14)$$

If there are no *external* forces, equation (1) gives

$$m_1 u_1 + m_2 u_2 + m_3 u_3 + \text{etc.} = \text{a constant.}$$

Hence $\bar{u}$ is a constant; that is, the average distance of the matter from the fixed plane increases at a constant rate. Of course, u_1, u_2, etc., may change, owing to the action of the different particles on each other; but these changes are such that $\bar{u}$ does not change.

If there are external forces, there will be changes in the velocities u_1, u_2, etc., owing to the external action in addition to those due to the mutual action of the bodies. Consequently $\bar{u}$ will also change. Let $\bar{a}$ be the rate of change of $\bar{u}$, a_1 that of u_1, a_2 that of u_2, etc. Hence, from (14) it follows that

$$\begin{aligned}(m_1 + m_2 + m_3 + \text{etc.})\,\bar{a} &= m_1 a_1 + m_2 a_2 + m_3 a_3 + \text{etc.}\\ &= X_1 + X_2 + X_3 + \text{etc.,}\end{aligned}$$

where X_1 is the entire force of m_1, in the direction away from the fixed plane, including both external and internal forces; X_2 applies similarly to m_2, etc. But the sum of all the *internal* forces in any fixed direction equals zero, because "action and reaction are equal and opposite." Hence $X_1 + X_2 + X_3 +$ etc. equals simply the sum of the external forces in the direction away from the fixed plane; and it may be written X. Consequently,

$$(m_1 + m_2 + m_3 + \text{etc.})\, \bar{a} = X, \quad . \quad . \quad . \quad (15)$$

which states that the acceleration away from the plane of the average distance of the system from the plane is equal to the total external force in that direction divided by the total mass.

39. We might in a similar way consider the average distance of the system (or the single body) from any other plane. In particular, let us take three fixed planes at right angles to each other, e. g. the three planes meeting in the corner of a room, and find the average distances of the systems from them, and the changes in these distances We may write

$$\left.\begin{aligned}(m_1 + m_2 + m_3 + \text{etc.})\, \bar{x} &= m_1 x_1 + m_2 x_2 + m_3 x_3 + \text{etc.}\\ (m_1 + m_2 + m_3 + \text{etc.})\, \bar{y} &= m_1 y_1 + m_2 y_2 + m_3 y_3 + \text{etc.}\\ (m_1 + m_2 + m_3 + \text{etc.})\, \bar{z} &= m_1 z_1 + m_2 z_2 + m_3 z_3 + \text{etc.}\end{aligned}\right\} (16)$$

where the x, y, z's refer to the distances from the three planes. But since $\bar{x}$, $\bar{y}$, $\bar{z}$, have definite numerical values at any instant, they determine the position of a *point* whose distance from one plane is $\bar{x}$, i. e. the average distance of the entire system from that plane, and similarly for the other planes. Consequently this *point* represents the average distance of the system from the three fixed planes; and it is called the "centre of inertia" of the system. It is not a point of the body, but a geometrical point in space.

Thus the centre of inertia of a uniform rod is its middle point. For, consider the rod made up of equal separate masses, and let m_1 and m_2 be two which are at the ends. Take as the plane of reference one perpendicular to the rod, and let x_1 and x_2 be the distances of m_1 and m_2 from the plane.

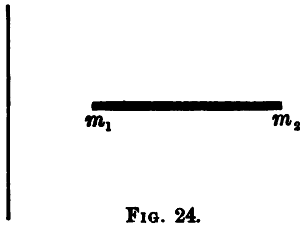


FIG. 24.

$$(m_1 + m_2)\,\bar{x} = m_1 x_1 + m_2 x_2.$$

But $m_1 = m_2$; hence $\bar{x} = \dfrac{x_1 + x_2}{2}$, i. e. the centre of inertia of these two masses is half-way between them. Similarly for the other masses which make up the rod, always combining those which are equidistant from the two ends. The centre of inertia of a uniform sphere (or spherical shell) is also its centre of figure. Other illustrations will be given later.

The velocity of the centre of inertia away from the three planes is given by the three equations

$$\left.\begin{aligned}(m_1 + m_2 + m_3 + \text{etc.})\,\bar{u} &= m_1 u_1 + m_2 u_2 + m_3 u_3 + \text{etc.}\\ (m_1 + m_2 + m_3 + \text{etc.})\,\bar{v} &= m_1 v_1 + m_2 v_2 + m_3 v_3 + \text{etc.}\\ (m_1 + m_2 + m_3 + \text{etc.})\,\bar{w} &= m_1 w_1 + m_2 w_2 + m_3 w_3 + \text{etc.}\end{aligned}\right\}\quad(17)$$

And the actual resultant velocity is found by adding $\bar{u}$, $\bar{v}$, $\bar{w}$, geometrically, so as to give the diagonal of a rectangular parallelopiped of which they are the edges.

If there are no external forces, $\bar{u}$, $\bar{v}$, $\bar{w}$ are constant; so, also, is their geometrical sum. Consequently, the centre of inertia will move in a straight line with a constant speed, or will remain at rest. As an illustration of this fact, it is evident that if any number of bodies are at rest, and are then set in motion by their mutual actions, they will so move that their centre of inertia always remains the same as it was before the motion. Thus, when a bullet is fired from a

rifle so suspended as to be free to move, the resulting motion will be such as to leave the centre of inertia of the rifle and the bullet unchanged. If a boiler explodes, the parts will so move that the centre of inertia does not change. If the solar system, i. e. the sun and planets with their satellites, is free from external action, its centre of inertia must be moving through space with a constant velocity, if it is in motion at all.

If there are external forces,

$$\left.\begin{aligned}(m_1 + m_2 + m_3 + \text{etc.})\,\bar{a} &= X\\(m_1 + m_2 + m_3 + \text{etc.})\,\bar{b} &= Y\\(m_1 + m_2 + m_3 + \text{etc.})\,\bar{c} &= Z\end{aligned}\right\} \quad . \quad . \quad . \quad (18)$$

where $\bar{a}, \bar{b}, \bar{c}$ are the accelerations of the centre of inertia in the three directions; and X, Y, Z are the sums of the *external* forces in those directions, because all the internal forces are balanced. The actual acceleration of the centre of inertia is the geometrical sum of $\bar{a}$, $\bar{b}$, $\bar{c}$.

X is the sum of the components in a certain direction of all the external forces "acting on" each one of the particles; similarly Y and Z refer to the components in the other two directions. If these components were to act at a single point, they would have a resultant effect equal to their geometrical sum. Consequently adding geometrically the three equations, and writing M for the entire mass, A for the actual acceleration of the centre of inertia, and R for the geometrical sum of X, Y, Z,

$$MA = R \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

This equation states that, when there are external forces, the motion of the centre of inertia of a system of bodies (or a single body) is exactly the same as would be that of a particle of mass, M, "acted upon by" a force, R, where M is the total mass of the system, and R is the geometrical sum of all the external forces. Or, so far as translation is concerned, the motion of any body is exactly the same as

would be that of a particle whose mass equals the entire mass, if "acted upon" by all the external forces directly.

Thus, consider a uniform stick thrown at random along a horizontal table. The centre of inertia of the stick is its middle point; and, as the stick moves, this point must follow a straight line, however the stick itself may revolve; because, if a body whose mass were equal to that of the stick were acted upon by a force equal to that used on the stick, it would move in a straight line.

Again, consider a bomb-shell fired in the air. If it does not burst, its centre of inertia, i. e. its centre of figure, will describe a certain path. But even if it does burst, the fragments will so move that their centre of inertia will follow identically the same path, because the *external* forces are not changed.

40. Special Cases of Centres of Inertia.

1. Triangular board, $A\,B\,C$.

Draw the three medial lines $A\,a$, $B\,b$, $C\,c$, connecting the vertices with the middle points of the opposite sides. They meet in a point O. Since the straight line $A\,a$ divides the triangle into two equal halves, the centre of inertia must lie on it; for the triangle may be considered built up of a great number of strips parallel to the side $B\,C$, and as the centre of inertia of each of these strips lies on the medial line $A\,a$, the centre of inertia of the entire triangle must lie on it also. Similarly, it must lie on $B\,b$ and $C\,c$; that is, it must be the point O, their common point of intersection.

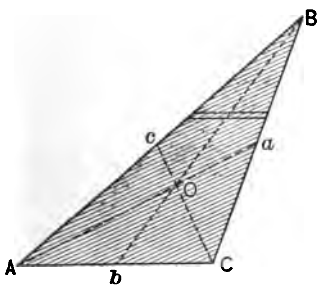


FIG. 25.

2. A uniform rod, mass $m_3 = 25$, carrying two symmetrical bobs whose masses are $m_1 = 15$, $m_2 = 20$; the dimensions and distances being as indicated (Fig. 26).

The centre of inertia of the rod itself is its middle point; that is, the rod acts as if its mass were concentrated at that point, which is at a distance 15 cm. from the ends. Take as a plane from which to measure distances one perpendicular to the rod at its left end. Then

$$m_1 = 15,\ x_1 = 5;\quad m_2 = 20,\ x_2 = 20;\quad m_3 = 25,\ x_3 = 15;$$

and $$\bar{x} = \frac{m_1x_1 + m_2x_2 + m_3x_3}{m_1 + m_2 + m_3} = \frac{75 + 400 + 375}{60} = 14.17.$$

The centre of inertia must, then, lie at a distance of 14.17 cm. from the plane at the end of the rod; and since the bobs are symmetrical, it must lie in the axis of the rod at that distance from the end.

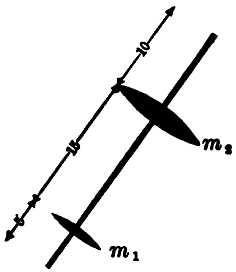


FIG. 26.

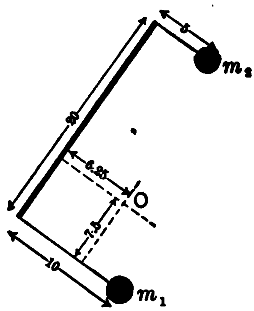


FIG. 27.

3. A rigid framework lying in a plane; two bodies, whose masses $m_1 = 20$, $m_2 = 10$, are connected by massless wires to a uniform rod whose mass $m_3 = 10$; the dimensions being as shown.

Take as the two planes of reference one perpendicular to the rod at its lower end, the other through the rod perpendicular to the two wires.

Hence
$$m_1 = 20,\quad x_1 = 0,\quad y_1 = 10;$$
$$m_2 = 10,\quad x_2 = 20,\quad y_2 = 5;$$
$$m_3 = 10,\quad x_3 = 10,\quad y_3 = 0.$$

So
$$\bar{x} = \tfrac{300}{40} = 7.5;\quad \bar{y} = \tfrac{250}{40} = 6.25.$$

That is, the centre of inertia is a point at a distance 7.5 cm. from the plane perpendicular to the rod at its lower end; and a distance 6.25 cm. from the rod itself in a direction parallel to the wires; therefore it is at the point O as shown.

Matter in Rotation

41. There is another general law which is obeyed by a system of bodies in motion, and which is not a necessary consequence of the law which has just been dis-

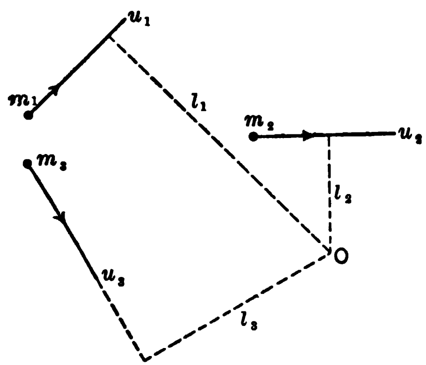


Fig. 28.

cussed, and which states that $m_1 u_1 + m_2 u_2 + m_3 u_3 +$ etc., = a constant if there is no external influence. This second law may be stated in this way: Let m_1, m_2, m_3, etc., be the masses of a system of particles; let their velocities in any plane (or in planes parallel to each other) be u_1, u_2, u_3, etc., as shown; let s_1, s_2, s_3, etc., be their speeds; and let O be the position of *any* fixed line perpendicular to these parallel planes. Draw perpendiculars from this line to the directions of the velocities, and let l_1, l_2, l_3, etc., be their lengths. Then the general law is that

$$m_1 s_1 l_1 + m_2 s_2 l_2 + m_3 s_3 l_3 + \text{etc.} = \text{a constant} \qquad (20)$$

if there are no external influences. In this sum, a product is called positive, if looking at m from O it is moving towards the right; thus, $m_1\, s_1\, l_1$ and $m_2\, s_2\, l_2$ are positive; $m_3\, s_3\, l_3$ is negative. The product $m\, s\, l$ is called the "moment of momentum" of the mass m moving with a speed, s, in a line at a distance, l, from the arbitrary axis; and this law is called the "Principle of the Conservation of Moment of Momentum."

This law is thought to be true in all cases, and thus to be one of the great laws of nature, because no exceptions to it are known; all observed phenomena are in aocord with it.

Special Cases of Rotation

42. 1. Let there be a single particle. Then $m\, s\, l =$ a constant, if there are no external influences. (This might be considered also as a consequence of the general law that "$m\, u =$ a constant, if there are no external influences.") But $m\, s\, l$ will still remain a constant when there are external influences, if these are of a definite kind. Thus, let a particle whose mass is m be connected to a pivot, O, by a rigid beam of length, r, whose mass is so small that it may be neglected. The body will then move in a circle around the pivot. Measuring l from this pivot, $l = r$, $s = r\, \omega$ $\therefore m\, s\, l = m\, r^2\, \omega$, where ω is the angular *velocity*, because the product $m\, s\, l$ is positive for motion around a definite axis in a definite direction. Now, it is an observed fact that ω will remain constant, and hence $m\, r^2\, \omega$ also, if all the forces acting on m have directions which pass through the pivot. Forces of this kind simply alter the pressure on the pivot.

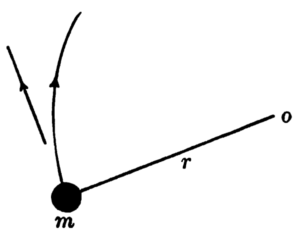


FIG. 29.

43. If there is a force "acting on" m, whose direction does not pass through the pivot, it is observed that there is a change in the angular speed; i. e. $m r^2 \omega$ changes. The rate of change of $m r^2 \omega$ is called the "moment around the axis" to which r and ω refer. $m r^2$ cannot change if r is constant; and the rate of change of ω is the angular acceleration. Call the moment, L; the angular acceleration, a; and mr^2, I. Then

$$L = I a \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (21)$$

This is perfectly analogous to the similar equation in motion of translation,

$$F = m\,a.$$

The moment of inertia, I, measures inertia of matter for rotation, as mass, m, does for translation. And it is seen to depend upon not alone the mass but also the distance of the particle from the axis of rotation. If there are several particles revolving around the same axis, the entire moment of inertia is the sum of the individual ones, just as the entire mass is the sum of the separate ones. Further, just as $m\,u$ is called the linear momentum, so $I\,\omega$ is called the "angular momentum." That is, in the case of motion in a circle around a fixed axis, the moment of momentum is called the angular momentum.

As explained in Chapter I. Article 16, angular accelerations may be compounded, if the axes meet in a point. Consequently moments may also be added, or resolved into components.

A moment is produced, as stated above, when the direction of the force does not pass through the pivot; and the connection beween moment and force is easily found.

Consider a particle whose mass is m, revolving in a circle of radius, r, around a pivot, O, with an angular velocity which at any instant equals ω. $I = m\,r^2$.

Hence $I\omega = m r^2 \omega = m r u$, where u is the linear velocity of m at that instant. Let a *force*, F, "act" perpendicular to the radius, $F = m a$; but the moment is the rate of change of $m r u$, that is, $L = m r \,.\, a$; hence

$$L = F r.$$

Therefore the moment around an axis equals the product of the force and the perpendicular distance from the

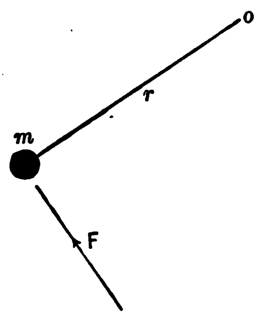


FIG. 30.

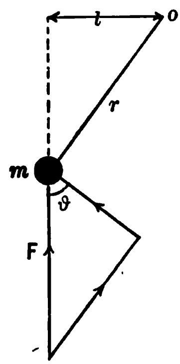


FIG. 31.

axis to the direction of the force. Similarly, if the force F is applied obliquely to the radius. The moment, by definition, is

$$L = m a \,.\, r.$$

The component of the force F perpendicular to the radius is $F \cos\theta$; and this must equal $m a$; i. e.

$$F \cos\theta = m a.$$

Hence $$L = F r \cos\theta = F l, \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

where l is the perpendicular distance from the axis to the direction of the force. It is not necessary to take into account the other component of the force F, because it is along the radius and so passes through the pivot, simply producing a pressure there.

44. 2. Let there be any number of particles which are rigidly connected, and which are at any instant turning around an axis. An illustration of this is an ordinary solid in rotation.

Let m_1, m_2, m_3, etc., be the masses; r_1, r_2, r_3, etc., their distances from the axis; and ω, the angular velocity common to all. Hence $l_1 = r_1$, $s_1 = r_1\omega$; $l_2 = r_2$, $s_2 = r_2\omega$; $l_3 = r_3$, $s_3 = r_3\omega$; etc.; and so

$$m_1l_1s_1 + m_2l_2s_2 + m_3l_3s_3 + \text{etc.} = (m_1r_1^2 + m_2r_2^2 + m_3r_3^2 + \text{etc.})\,\omega.$$

$m_1r_1^2 + m_2r_2^2 + m_3r_3^2 +$ etc. is the sum of the moments of inertia of m_1, m_2, m_3, etc., around the axis of rotation; call it the total moment of inertia of the system around the axis, and write it I. The law is that the product

$$I\omega = \text{a constant} \quad . \quad . \quad . \quad . \quad . \quad (23)$$

if there are no external influences. This is verified by experiment.

Thus, if a fly-wheel is set spinning around an axle, it would continue revolving with a constant speed forever were it not for friction.

But if there are external influences, $I\omega$ can be altered in two ways, — by the direction of the axis changing, and by the angular speed changing. To produce an angular speed around any axis, a moment around that axis is necessary, as was shown in Case 1. And, as was proved in Chapter I. (Art. 24), if this additional angular speed is around the same axis as the existing motion, the angular speed will be changed, but not the direction of the axis; if the additional angular speed is around an axis perpendicular to that of the existing motion, then the direction of the axis is changed. Consequently, the product $I\omega$ will remain a constant unless the external forces, if there are any, produce a suitable moment.

45. Illustrations. *a.* A disc rotating around a fixed axis. Let the disc be pivoted at the point O, and let there be a force, F, in the plane of the disc applied at the point P (e. g. let a string which is fastened to a nail at P be pulled with a force, F).

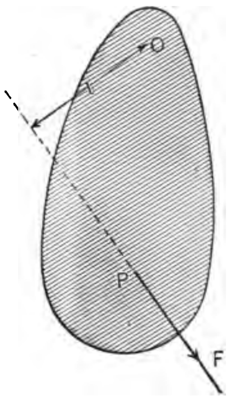


FIG. 32.

Since the disc is a rigid body, the effect of the force F is the same when applied at P as if applied at any point in the line of the force; and its moment around the axis O is Fl, where l is the perpendicular distance from O to the line of action of F. Then, if I is the moment of inertia about the pivot,

$$Fl = Ia \quad . \quad . \quad . \quad (24)$$

This moment in the case as shown produces angular acceleration in a direction opposite to the motion of the hands of a watch. Let there be another force whose moment around the axis O is equal to the former moment, but which produces acceleration in the same direction as the motion of the hands of a watch; then, if both these moments "act on" the body, they will neutralize each other, and the angular acceleration will be zero; that is, the angular velocity will remain constant.

46. *b.* Two discs which are so arranged as to rotate around a fixed axis. After being set in rotation, let them be pushed together. Then, owing to friction between their surfaces, they will gradually assume the same angular speed. The two discs form a system entirely removed from external influences, except the push which

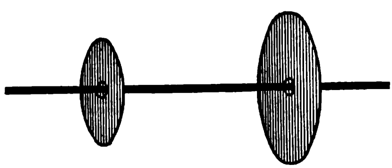


FIG. 33.

brings them together. This, however, is a force along the axis, and so produces no moment. Consequently, according to the general law, the total angular momentum must remain unchanged. Let I_1 and I_2 be the moments of inertia of the two discs around the axis, ω_1 and ω_2 their angular velocities before coming in contact, ω_3 the angular velocity common to both after the speed becomes the same for both. Then it follows that

$$I_1 \omega_1 + I_2 \omega_2 = (I_1 + I)\, \omega_3 \quad . \quad . \quad . \quad . \quad (25)$$

This is perfectly in accord with experiment.

The analogy should be noticed between this phenomenon and that of the impact of two bodies which stick together after meeting. In this latter

$$m_1 u_1 + m_2 u_2 = (m_1 + m_2)\, u_3.$$

47. *c.* A cylinder *rolling* down an inclined plane.

The moment in this case is due to the earth's force, $m\,g$, and the axis of rotation is at any instant the line of con-

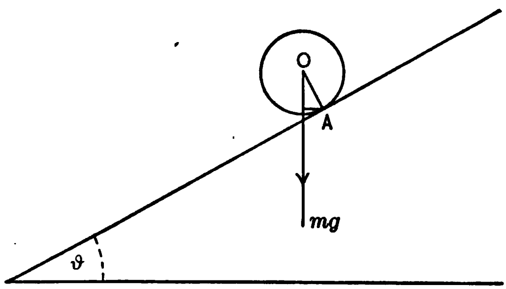


FIG. 34.

tact between the cylinder and the plane. The force acts at the centre of inertia of the cylinder; and consequently, calling r the radius of the cylinder, the moment of the force about the axis of rotation is the product of $m\,g$ and the perpendicular distance from A to a vertical line through the centre of the cylinder. That is,

$$L = m g r \sin\theta.$$

And so $$m g r \sin\theta = I a.$$

It may be proved by higher mathematics that in this case

$$I = \tfrac{3}{2} m r^2.$$

Hence $$a = \tfrac{2}{3} g \frac{\sin\theta}{r}.$$

Thus a is a constant; and as the direction of the axis does not change, the resulting angular speed and angle turned through may be calculated from Formulæ (10), (11), and (12), Chapter I. Article 22.

48. *d*. A spinning top whose axis is not vertical. Owing to the action of the earth, there is a moment tending to make the top turn over; that is, the moment is around an axis perpendicular to the axis of figure. If the top is spinning, the motion will then be compounded of the velocity around the axis of figure and an acceleration around an axis perpendicular to it. Hence the motion is as described in Chapter I. Article 23.

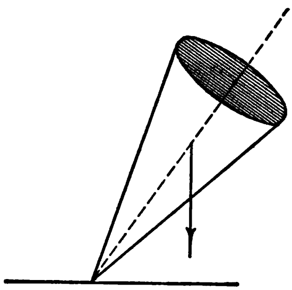


Fig. 35.

49. Centre of Inertia. If the only force acting on a system of rigidly connected particles is applied in such a direction as to pass through the centre of inertia, the system will not be set in rotation (or, if it has any angular velocity, this will not be altered). A special case, of course, of such a system is a single solid body.

This theorem may be proved in this way: Let F be a force whose direction passes through the centre of inertia of the system; m_1, m_2, m_3, etc., the masses of the different particles composing the rigid system; x_1, x_2, x_3, etc.,

their distances from a fixed plane which is parallel to the direction of the force; a_1, a_2, a_3, etc., their linear accelerations parallel to the direction of the force; and O the

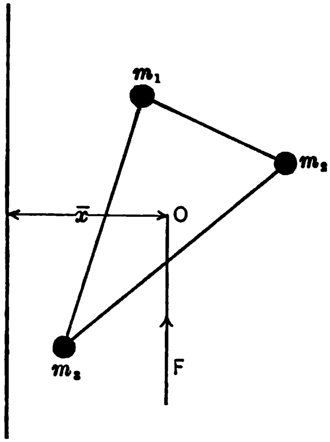


FIG. 36.

centre of inertia at a distance, $\bar{x}$, from the plane. Then, from (16) and (18),

$$\bar{x} = \frac{m_1 x_1 + m_2 x_2 + m_3 x_3 + \text{etc.}}{m_1 + m_2 + m_3 + \text{etc.}}$$

$$\bar{a} = \frac{m_1 a_1 + m_2 a_2 + m_3 a_3 + \text{etc.}}{m_1 + m_2 + m_3 + \text{etc.}}.$$

The sum of the moments of the forces acting on each particle, taken around any axis which is in the fixed plane, and which is perpendicular to the direction of the force, is

$$m_1 a_1 x_1 + m_2 a_2 x_2 + m_3 a_3 x_3 + \text{etc.}$$

and this moment of the system must equal the moment of the force around the same axis, $F\bar{x}$. From (19)

$$F = (m_1 + m_2 + m_3 + \text{etc.})\,\bar{a}.$$

Hence

$$(m_1 + m_2 + m_3 + \text{etc.})\,\bar{a}\,\bar{x} = m_1 a_1 x_1 + m_2 a_2 x_2 + m_3 a_3 x_3 + \text{etc.}$$

That is, substituting for $\bar{x}$ its value,

$$m_1 \bar{a} x_1 + m_2 \bar{a} x_2 + m_3 \bar{a} x_3 + \text{etc.} = m_1 a_1 x_1 + m_2 a_2 x_2 + m_3 a_3 x_3 + \text{etc.}$$

And as this same equation must hold for different values of the x's (because the plane is *any* fixed plane parallel to the force), their coefficients on the two sides of the equation must be identical; that is,

$$\bar{a} = a_1 = a_2 = a_3 = \text{etc.}, \quad . \quad . \quad . \quad . \qquad (26)$$

which means that all the portions of the system have the same linear acceleration, so that there is no rotation produced.

50. If the force is applied in a direction which does not pass through the centre of inertia, there will be both translation and rotation; the centre of inertia will move as if the entire mass were concentrated there and "acted upon" by the total force, and the body will itself rotate around the centre of inertia as it moves. It may be proved that the rotation produced around axes passing through the centre of inertia as it is moving is exactly the same as it would be if the centre of inertia was fixed.

Harmonic Motion

51. An illustration of harmonic motion is given by a simple pendulum, which consists of a particle of mass, m, suspended from a fixed point, P, by a string of length, l, whose mass is too small to be taken into account. If the pendulum is set in vibrations in a vertical plane, there will be a moment due to the earth's action which will always tend to bring the pendulum back into its vertical position. As was stated in Article 33, there is a force, $m\,g$, acting vertically down on m; and hence, when the angular displacement of the pendulum is θ, the moment around the axis P

is $m\,g\,l \sin \theta$. The moment of inertia of m around P is $m\,l^2$. Hence, since the angular acceleration is the ratio of the moment of the force to the moment of inertia (Art. 43), its value is

$$\frac{m\,g\,l \sin \theta}{m\,l^2} = \frac{g}{l} \sin \theta.$$

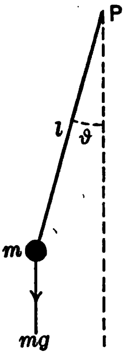


FIG. 37.

But if the angle is *small*, $\sin \theta = \theta$. So the angular acceleration is $\frac{g}{l}\,\theta$; and since it opposes displacement from the vertical position, the vibrations are harmonic (Art. 25). The period, then, for *small* vibrations is

$$T = 2\,\pi \sqrt{\frac{l}{g}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (27)$$

It is possible that g might be different for different masses or for different kinds of matter; but this is proved by direct experiment not to be true.

Consequently a pendulum of any kind may be used to measure equal intervals of time, since its period is a constant, depending only on its length.

The "second's pendulum" is a pendulum which "beats" seconds; that is, its complete period is 2 seconds. Hence, $T = 2$; and, if $g = 980$,

$$l = \frac{T^2 g}{4\,\pi^2} = \frac{980}{\pi^2} = 99.3 \text{ cm.}$$

A pendulum, then, of this length, swinging at a place where $g = 980$, would make a half vibration each second.

The motion of a simple pendulum may be regarded also as a case of translation; and, if the vibration is small enough, the arc is practically a straight line. The *force* which tends to bring the pendulum back into its vertical position is $m\,g \sin \theta$; so the linear acceleration is $g \sin \theta$

or $g\,\theta$, if the arc is small. The linear displacement of the pendulum bob from its vertical position is $l\,\theta$. Call this x. Then

linear acceleration $\qquad a = g\,\theta$

linear displacement $\qquad x = l\,\theta$

$$\therefore \qquad a = \frac{g}{l}\,x.$$

And as the acceleration is always towards the central position, the motion is harmonic (see Art. 21); and the period of the small vibrations is

$$T = 2\,\pi\sqrt{\frac{l}{g}}.$$

The tension in the string does not influence the motion in the least, as its line of action passes through the pivot, and so its moment is zero.

Motion in General of a Rigid Body

52. It has been proved (Art. 36) that, if a body is so small that it can be considered a particle, its motion under the action of two forces, represented by the lines $\overline{O\,A_1}$ and $\overline{O\,A_2}$, is given by their sum, a force represented by $\overline{O\,B}$. Further, in the case of a large body, it has been proved that, so far as translation is concerned, the motion of the centre of inertia is just what the motion of a particle would be if its mass were that of the body, and if all the forces acted on it directly. The motion of rotation of a large body is also known, if there is but a

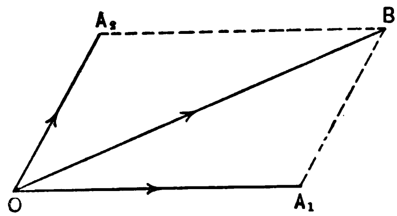


Fig. 38.

single force acting on it. The problem now is to find, not alone the motion of translation, but also that of rotation, when a large body is acted on by any number of forces.

This general problem is quite difficult; and so only a special case will be considered here, viz. that in which the directions of the forces are lines in one plane. Such forces are called "co-planar." The first step will be to find, if possible, a single force which will in every way produce the same effect as all the forces acting together. Such a force which is equivalent to the combined action of all the individual forces is called their "resultant."

The resultant of two forces must, then, produce the same translation as these two combined, and must also have the same moment around any axis as the combined moments of the two forces. If the resultant of two forces can be found, it can be combined with a third force, etc.; and so the resultant of any number of forces can be found.

It will be necessary to distinguish two cases: *a*, two non-parallel forces; *b*, two parallel forces.

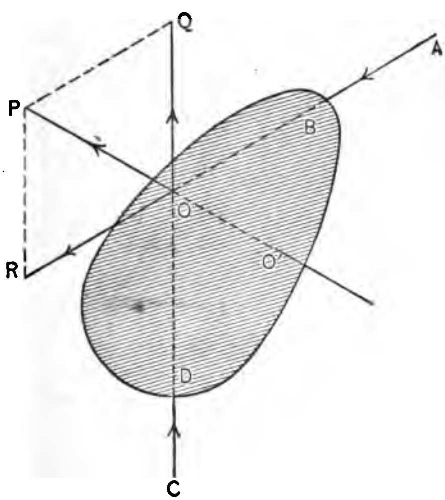


FIG. 39.

53. Non-parallel Forces. Two non-parallel forces lying in the same plane and represented by the lines $\overline{AB}$ and $\overline{CD}$, act on a rigid body at the points B and D, respectively (e. g. two strings are fastened to nails at B and D, and are then pulled in the given directions with the specified forces). Prolong the lines of

action until they meet in a point, O. Let it be in the body itself. Since the body is rigid, the force $A\,B$ will produce the same effect if applied at any point in its line of action, e. g. at the point O. Similarly the force $C\,D$ may be considered as also applied at O. Since the two forces may be regarded as acting at the point O, the motion of translation of O must obviously be given by the geometrical sum of the forces taken at that point. That is, lay off from O two lines, $\overline{O\,R}$ and $\overline{O\,Q}$, equal in length and direction to $\overline{A\,B}$ and $\overline{C\,D}$, and add them. Their sum, $\overline{O\,P}$, is then equivalent to them so far as translation is concerned, and may be considered applied at any point in its line of action, e. g. O or O'.

But the force $\overline{O\,P}$ will also have the same moment around any axis as the combined moments of its two components, $\overline{A\,B}$ and $\overline{C\,D}$. This may be proved in the following way: Take *any* axis perpendicular to the plane of the forces, and let its intersection with the plane be the point S. The moment of $\overline{O\,Q}$ around S equals the product of $\overline{O\,Q}$ and the perpendicular distance from S to the direction $O\,Q$, i. e. $\overline{S\,S_1}$; but this is equal to a number which is twice the area of the triangle $(O\,S\,Q)$. Similarly, the moment of $\overline{O\,R}$ around S equals a number which is twice the area of the triangle $(O\,S\,R)$; but, in accordance with the definition of Article 41, this must be considered negative. Hence the combined moment of $\overline{O\,Q}$ and $\overline{O\,R}$ around S is equal to twice the dif-

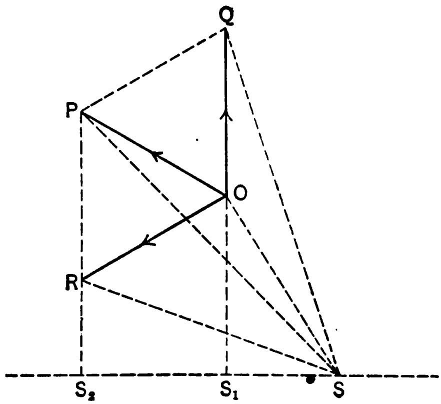


Fig. 40.

ference in the areas of the triangles $(O\,S\,Q)$ and $(O\,S\,R)$. But, by ordinary geometry,

$$(O\,S\,Q) = (R\,S\,P) - (P\,O\,R)$$

Hence $$(O\,S\,Q) - (O\,S\,R) =$$
$$(R\,S\,P) - [(P\,O\,R) + (O\,S\,R)] = -(O\,S\,P).$$

But $-2\,(O\,S\,P)$ is the moment of $\overline{O\,P}$ around S. Consequently, the moment of $\overline{O\,P}$ equals the sum of the moments of $\overline{O\,Q}$ and $\overline{O\,R}$, around any axis. Therefore the force $\overline{O\,P}$ is the resultant of $\overline{O\,Q}$ and $\overline{O\,R}$; that is, of $\overline{A\,B}$ and $\overline{C\,D}$.

Consequently, the resultant of two non-parallel forces which lie in the same plane is a force which also lies in this plane, whose numerical value and direction is given by the geometrical sum of the two forces, and whose line of action is such as to pass through the point of intersection of the two forces.

The same demonstration holds even when O, the point of intersection of the two forces, is not a point of the body. For, as has just been proved, when the point is in the body, a force, $\overline{O\,P}$, the geometrical sum of the two forces, is the resultant; and it may be applied at any point, O', in the line of action of $\overline{O\,P}$. But, now let the portion of the body surrounding the point O be removed; no change is made in the two forces, the resultant is the same; and yet the point of intersection is not now a point of the body. As an illustration, let two forces, F_1 and F_2, act on a rigid bar. Prolong the lines of action

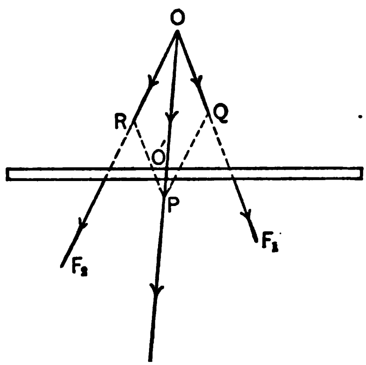


FIG. 41.

until they meet at O; lay off $\overline{OQ} = F_1$, $\overline{OR} = F_2$; construct their sum, $\overline{OP}$. Then the resultant of F_1 and F_2 is a force equal to $\overline{OP}$, applied at the point, O', where the direction of the force meets the bar.

54. Parallel Forces. Two parallel forces, F_1 and F_2, act on a rigid body at the points A and B. respectively. To

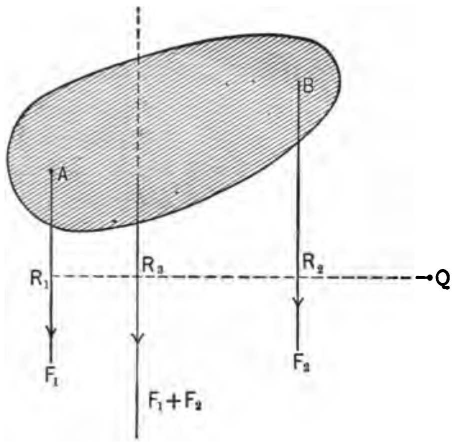


FIG. 42.

FIG. 43.

find their resultant, consider first the case where the two forces are not quite parallel: the lines of action will meet in a point, O, at a great distance. Lay off from this point, O, two lines equal to F_1 and F_2; their geometrical sum is the line $\overline{OP}$. But, in the limit, when the lines are parallel, the line $\overline{OP}$ equals the algebraic sum $F_1 + F_2$, and is parallel to the component forces. Hence the resultant of two parallel forces lies in their plane, is equal to their algebraic sum, and is parallel to them. Further, it must occupy such a position with relation to F_1 and F_2 that its moment around *any* axis perpendicular to the plane is equal to the sum of the moments of

F_1 and F_2 around the same axis. Let Q be the intersection of any such axis with the plane; draw a perpendicular from Q to the three parallel lines representing the forces, intersecting them in the points R_1, R_2, R_3. Then the moment of F_1 around the axis is $F_1 \times \overline{R_1 Q}$; that of F_2, $F_2 \times \overline{R_2 Q}$; that of $F_1 + F_2$, $(F_1 + F_2) \times \overline{R_3 Q}$. Hence $\overline{R_3 Q}$ must have such a length that

$$F_1 \overline{R_1 Q} + F_2 \overline{R_2 Q} = (F_1 + F_2)\, \overline{R_3 Q}.$$

Hence $$F_1 \overline{R_1 R_3} + F_2 \overline{R_2 R_3} = 0$$

or $$F_1 \overline{R_1 R_3} = F_2 \overline{R_3 R_2} \quad . \quad . \quad . \quad . \quad (28)$$

That is, R_3 is such a point that the moment of F_1 around an axis through it perpendicular to the plane equals the moment of F_2 around it. So the relative position as well as the numerical value and direction of the resultant is fixed. It may be applied at any point in its line of action.

55. As a special case, consider two particles whose masses are m_1 and m_2, and which are connected by a massless beam. Let them be under the action of the earth; the mutual influence of the two particles is zero (Art. 32), but there is a force $m_1 g$ acting on m_1, and a parallel force $m_2 g$ acting on m_2. The resultant, then, is $(m_1 + m_2)\, g$ in a parallel direction, and so placed that

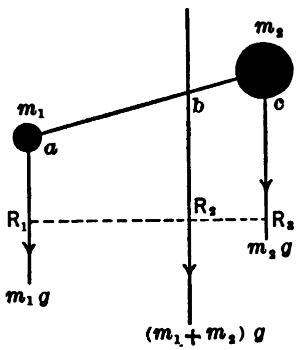


FIG. 44.

$$m_1 g \overline{R_1 R_3} = m_2 g \overline{R_2 R_3}.$$

Its point of application will be where the line intersects the beam. By similar triangles this point, b, must be such that

$$m_1 \overline{a b} = m_2 \overline{b c}.$$

Hence b must be the centre of inertia of the two masses; because, taking as the plane of reference one perpendicular to the beam at the point a, the position of the centre of inertia, $\bar{x}$, is given by the equation,

$$(m_1 + m_2)\,\bar{x} = m_1 \cdot 0 + m_2\,\overline{a\,c}.$$

Hence, calling a point b' the centre of inertia,

$$\bar{x} = \overline{a\,b'}, \text{ and } (m_1 + m^2)\,\overline{a\,b'} = m_2\,\overline{a\,c}.$$

Thus $$m_1\,\overline{a\,b'} = m_2\,(\overline{a\,c} - \overline{a\,b'}) = m_2\,\overline{b'\,c}.$$

Therefore b' coincides with b. Similarly, if there are any number of particles (or a single large body) under the action of the earth, the resultant is the entire mass multiplied by g; and its line of action is vertically downward, and passes through the centre of inertia.

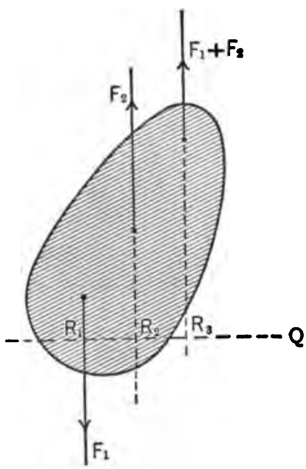


FIG. 45.

56. The same demonstration as given above in Article 54 holds good when the forces, though parallel, are in opposite directions. The resultant of the two forces, F_1 and F_2, is their *algebraic* sum, — that is, the difference between them, because they are in opposite directions; and it must be so placed that its moment around any axis equals the sum of the moments of the two components around the same axis. That is, using F_1 and F_2 as mere numbers,

$$F_2\,\overline{R_2\,Q} - F_1\,\overline{R_1\,Q} = (F_2 - F_1)\,\overline{R_3\,Q}.$$

Hence $$F_1\,\overline{R_1\,R_3} = F_2\,\overline{R_2\,R_2}.$$

57. Couples. Since the resultant, when the two parallel forces are in opposite directions, is $F_1 - F_2$; if the forces have equal numerical values, the resultant is zero. Such a combination of two equal parallel forces in opposite directions is called a "couple;" and there is no single force which is equivalent to it. The "strength" of the couple is a name given to the product of the numerical value of either of the forces and the perpendicular distance between them. The moment of the couple around *any* axis perpendicular to the plane of the forces is

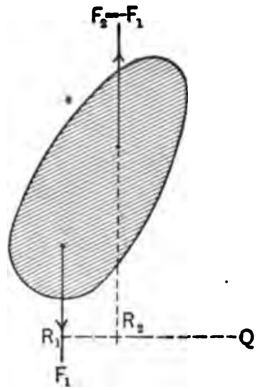


FIG. 46.

$$F_1\,\overline{R_1\,Q} - F_1\,\overline{R_2\,Q} = F_1\,\overline{R_1\,R_2};$$

that is, it equals the strength of the couple.

58. It may be proved without much difficulty that in the most general possible case of any number of forces, in one plane or not, acting on a rigid body, the resulting action is a single force combined with a single couple in a plane perpendicular to the force.

As was stated in Article 50, when a *single* force acts on a body, if its line of action passes through the centre of inertia, there is only translation; but if its line of action does not pass through the centre, there will be translation and also rotation around the centre of inertia as if it were a fixed point. Consequently, since, when any number of forces lying in one plane act on a body, there is a single resultant force (except in the case of two equal and opposite parallel forces), there is only translation if the resultant passes through the centre of inertia; otherwise there is both translation and rotation. If there is a couple, there is no translation, only rotation.

EQUILIBRIUM

59. A body or a system is said to be in "equilibrium" when there is no change in its existing motion; that is, when there is no change in its linear or angular momentum. The mathematical conditions are, therefore, if there are any external forces, (1) the sum of the components of the forces in any direction must equal zero; (2) the sum of the moments of the forces around any axis must equal zero.

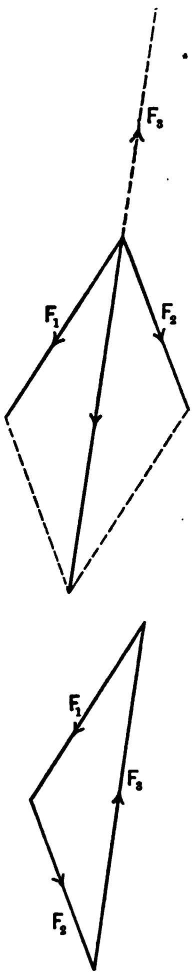


FIG. 47.

Special Cases. 60. 1. A particle acted upon by three forces. It was proved (Art. 52) that, when two forces, F_1 and F_2, acted upon a particle, their resultant was their geometrical sum. Therefore, if there is a third force, F_3, acting on the particle, which is equal but opposite to the resultant, it will exactly neutralize the action of F_1 and F_2. Consequently the particle will be in equilibrium. The geometrical condition, then, for the equilibrium of a particle under the action of three forces, F_1, F_2, F_3, is that they shall form a *closed* triangle when added geometrically.

Another and better way of considering this case is to form the components of the forces in any direction, and to express the fact that the sum must be zero.

As an illustration of the method, consider the equilibrium of a system as shown. $B\,C$ is a rigid beam which is held pressed perpendicularly against a vertical wall under the action of a string, $A\,B$, fastened to

the wall at A and to the beam at B, and a body of mass, m, which hangs vertically below B. (The string does not slip over B, but is fastened there.) The point B will come

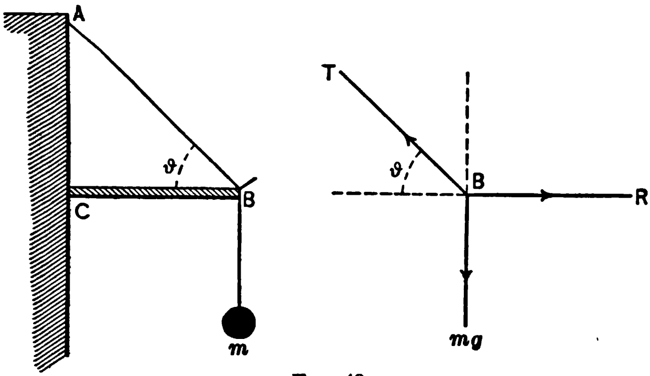


FIG. 48.

to rest, and is therefore in equilibrium; and three forces are acting there: (1) the force $m g$, due to the earth, vertically down; (2) the tension T of the string in the direction B to A; (3) the reaction R of the wall horizontally outward. (The beam is pressed against the wall, therefore there is an equal reaction of the wall against the beam.) Call the angle $A\,B\,C$, θ. Then the sum of the components, vertical and horizontal, are

$$T \sin\theta - m g \quad \text{and} \quad T\cos\theta - R,$$

and these must separately equal 0; i. e.

$$T\sin\theta - m g = 0,$$
$$T\cos\theta - R = 0.$$

These equations may be verified by direct experiment, as all the quantities can be measured separately.

61. 2. A large rigid body acted upon by three forces. It was proved in Articles 53 and 54, that when two forces acted on a body they had a resultant if they were in the

same plane and if they did not form a couple. Consequently, if the third force is equal and opposite to the resultant of the other two, and if it is applied in the same line as their resultant, there will be equilibrium. The two cases of non-parallel and parallel forces must be distinguished.

62. *a. Three Non-parallel Forces.* The resultant of two forces, F_1 and F_2, lying in one plane is found, as explained in Article 53, by prolonging their lines of action until they meet, and constructing at that point their geometrical sum. A force, F_3, then, equal and opposite to the resultant, must lie in the same plane as F_1 and F_2; its line of action must pass through the point of intersection of the lines of action of F_1 and F_2; and its numerical value must be the same as that of the resultant of F_1 and F_2, and it must be in the opposite direction. Hence, if a body is in equilibrium under the action of three non-parallel forces, they must all lie in the same plane, their lines of action must meet in the same point, the sum of their components in any direction must equal zero.

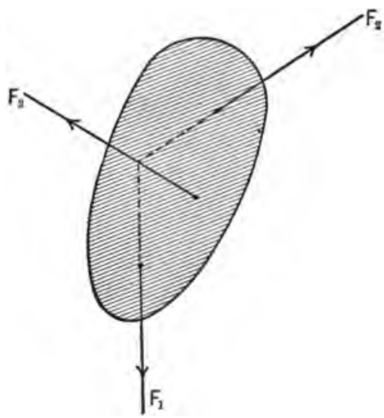


FIG. 49.

As an illustration, consider the case of a beam, $A\,B$, at rest lying on two *smooth* inclined planes. There are three forces acting on the beam: (1) the force $m\,g$ vertically down through the centre of inertia; (2) the reaction R_1 of the inclined plane at A, which is perpendicular to the plane since it is smooth; (3) the reaction R_2 of the second inclined plane at B. These forces must all lie in a plane, meet in

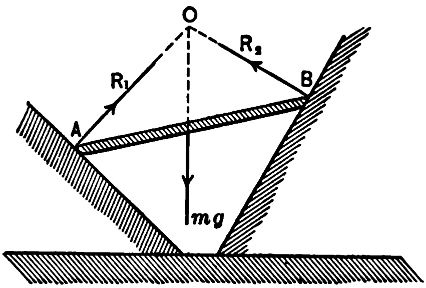


FIG. 50.

a point O, and neutralize one another, since the beam is at rest.

63. *b. Three Parallel Forces.* As explained in Articles 54 and 55, if there is a resultant of two parellel forces, it is a third force which lies in their plane, which is parallel to them, and whose numerical value equals their algebraic sum; and it is so placed relatively to them that its moment around any axis perpendicular to their plane equals the sum of their moments around the same axis. Therefore, if three parallel forces, F_1, F_2, F_3, are in equilibrium, the algebraic sum $F_1 + F_2 + F_3$ must equal zero, and the sum of the moments around any axis must equal zero.

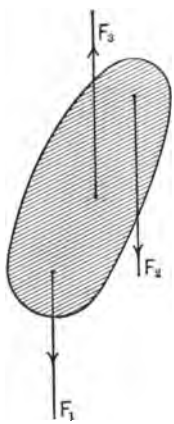


FIG. 51.

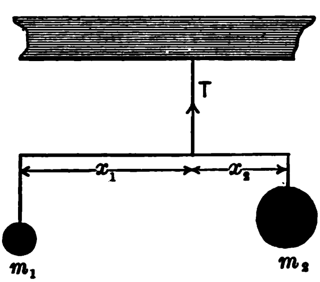


FIG. 52.

As an illustration, consider a rigid beam which carries two bodies whose masses are m_1 and m_2, and which is itself suspended by a string so placed that the system is at rest. (The beam is not necessarily horizontal.) Let the mass of the beam be so small that it may be

neglected. There are, then, three parallel forces acting on the beam: T, the tension of the string up; and $m_1 g$ and $m_2 g$ down. They are all in the same plane; and since the system is at rest,

$$T - m_1 g - m_2 g = 0.$$

Further, the points of application must be such that, taking moments around *any* axis perpendicular to the plane, the sum will be zero. Take them around an axis perpendicular to the plane of the forces at the point where the tension T is applied. The moment of T is zero; hence

$$m_1 g x_1 - m_2 g x_2 = 0.$$

These two conditions may be verified by experiment.

64. 3. A body acted upon by two couples. It was proved, in Article 57, that a couple composed of two parallel forces, each equal to F, whose distance apart was l, i. e. of strength $F l$, produced a moment $F l$ around any axis perpendicular to their plane. A couple has no resultant; and consequently its action can only be neutralized by another couple of equal strength but opposite to it. Thus, if two couples whose strengths are $F_1 l_1$ and $F_2 l_2$ act on a body, and if it is in equilibrium, they must lie in the same plane, and their strengths must have equal numerical values; but the moments must be in opposite directions. In the diagram the couple $F_1 l_1$ tends to produce rotation in

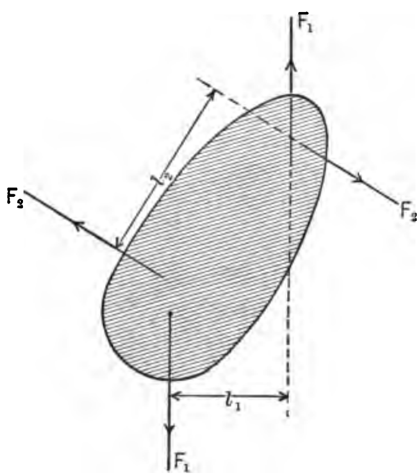


Fig. 53.

the direction opposite to the hands of a watch, while $F_2 l_2$ tends to produce rotation in the same direction as that of the hands. Consequently, if the body is in equilibrium, the numerical values are equal, i. e.

$$F_1 l_1 = F_2 l_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

65. Stability of Equilibrium. Even though two bodies or systems are in equilibrium, there may be differences in their conditions. Thus a cone balanced on its point is in equilibrium, as it is also when it is lying on its side or resting on its base; but there are obvious differences between the three states. Similarly an ellipsoid set in rotation around its longest axis may be in equilibrium, as it may be also when in rotation around its axis of intermediate length; but its properties are quite different in the two cases. The simplest way of considering these points of difference is to note what happens when the body or system is given a slight displacement from its condition of equilibrium.

Equilibrium is said to be "stable," if when the displacement occurs, the change produced becomes immediately less and less, and is finally reversed; so that the body or system returns towards its previous condition. In general, the system will then oscillate through the position of equilibrium. Thus, an ordinary pendulum when hanging at rest is in stable equilibrium; because, when it is struck a slight blow, the resulting velocity becomes less, and is finally reversed; and then the pendulum makes oscillations. Similarly, an ellipsoid spinning about its greatest or least axis is in stable equilibrium. These are illustrations of what are respectively called "statical" and "kinetic" stability.

Equilibrium is said to be "unstable," if, when the system is displaced, the initial change goes on increasing. Thus, a cone balanced on its apex is unstable, because, if it is given a blow, the initial velocity will increase. Similarly,

an ellipsoid spinning around its intermediate axis is unstable, because, if displaced, it will tend to spin around its greatest or its least axis.

Equilibrium is said to be "neutral," if, when the system is displaced, the initial change becomes permanent, and neither increases nor decreases. A sphere resting on a smooth table is in neutral equilibrium; because, if it is given a blow, the velocity thus produced will not be changed; and the sphere will move in a straight line with a constant speed.

66. Principle of Stable Equilibrium. The most common condition of equilibrium in nature is stability; because, of course, slight changes are always occurring; and so all bodies tend to pass into stable conditions. Any disturbance of stability must produce a reaction which tends to restore the body or system to its previous condition; and this principle can be applied to any stable condition, whether it is a purely mechanical one or not. Consider some illustrations of stability. 1. A body hanging suspended by a spiral spring is in stable equilibrium. If a blow downward is given it, the initial velocity will be decreased, owing to the increased tension of the spring. Hence, if the tension of a stretched spiral spring is increased by any means, it will raise the suspended body. 2. An iron bar surrounded by some medium, e. g. water, at a constant temperature is in stable equilibrium; for if its temperature is suddenly increased in any way, the tendency will be for it to return to the temperature of the surrounding medium. Now, when the temperature of an iron bar is increased, its length is increased; but this act of increasing in length produces a tendency for the bar to return to its former temperature. That is, if an iron bar is stretched by mechanical means, its temperature will fall. 3. Just the opposite effect happens with a piece of rubber cord. When its temperature is raised, it shrinks; consequently compressing the rubber diminishes its temperature, and stretching it raises its temperature.

Energy

67. Work and Energy. The words "to do work" convey a more or less definite idea to every one, and any body that has the power to do work is said to have "energy." Energy may be defined, then, as a measure of the possibility of doing work. Consider a few familiar illustrations. If a body is raised vertically from the earth, work is done by whatever raises it; if, for instance, a compressed spiral spring is allowed to expand and thus raise the body, the spring had energy when it was compressed and loses it when it is expanded. Work is done whenever the speed of a body is increased; thus, if a bullet is fired from a rifle, the powder has energy before it is exploded, and then loses it in the act of giving the bullet its motion. Work is required to make the fly-wheel of a steam-engine revolve faster; the steam in the cylinder has the energy, and then loses it. Work is done when a clock-spring is wound up, when a spiral spring is compressed, or when a gas is condensed. It is to be particularly noted that whenever work has been done upon any body, it itself is given the power of doing work; that is, it itself has energy given it as the result of another body losing energy. Thus, the body raised vertically above the earth can do work when it falls; for instance, it may compress the spiral spring again; the bullet in motion can also compress a spring, thus doing work and losing energy; the fly-wheel in rotation has energy, which can be used in winding up a spring or in working a pump; a compressed gas has energy also, because it can do work by expanding. Consequently it may be said, in general, that work is done as a result of one body losing energy and another gaining it; and experiments most carefully performed seem to prove definitely that *the amount of energy lost by one body exactly equals that gained by the second.* This means that starting with a certain amount of energy it is impossible by any

machinery or process to get more than a definite amount of work done by any one transformation.

Work is done, of course, in many other ways than those mentioned in the above paragraph. Thus, work is done when a piece of iron is pulled away from a fixed magnet; and if the iron is allowed to move back to the magnet, work may be done by it. Again, work is done whenever one portion of matter is made to move over another, as two boards rubbed together, a jet of water discharged into water or air, etc. And, as a result of this so-called "frictional" work being done, it is noticed that the temperature of the bodies rises, or some other so-called "effect of heat" is produced. It is possible, though, by making use of differences of temperature, to do work, as is shown by a steam-engine; and consequently, in this case too, energy is given the bodies as a result of the work done. It will be shown later on that all the "effects of heat" are in reality due to changes in the energy of the smallest portions of the matter in a body, not the energy of the body as a whole. A burning gas has energy, because it can do work; so has a Leyden jar "charged" with electricity; so has a Voltaic cell which produces an electric current; and so on indefinitely.

But to all these manifestations of energy the same law, as above stated, applies: the energy lost by one body or system is exactly equal to that gained by another. This principle has been confirmed by experiment in numerous ways; every observed phenomenon in nature is in perfect accord with it; and so it is regarded as one of nature's great laws. It is sometimes called the "Principle of the Conservation of Energy."

68. Measure of Work and Energy. It should be noticed that, in every case of work being done, only two ideas are involved: motion of a certain amount, and what has been called a force. When a body is raised vertically from the earth, there is a force downward "acting on" it

equal to $m g$ where m is its mass and g is a certain constant, as is proved by experiment (Art. 33); and the amount of work done depends upon the numerical value of this force and upon the height though which the body is raised. If a spiral spring is compressed, there is always a force acting in the opposite direction; and the amount of work done depends upon this force and the distance through which the spring is compressed. It is a question of distance, not time; because, for instance, when a spiral spring is compressed, the amount of energy it has cannot depend upon the time taken to compress it. And if there is no motion, no work is done. A spring kept compressed does no work; nor does a pillar supporting a building. When a body has its speed changed, there is by definition a force; and the longer the distance through which the force acts, the greater is the change.

The motion upon which depends the amount of work done is the distance traversed in the *direction* of the force, because, if a body moves at right angles to a force, no work is done. Thus, if a ball rolls along a horizontal table, no work is done under the action of the earth or against it. And if a body is moved in a direction, $\overline{A B}$, which is inclined to that of the force F, the work done depends on the distance $\overline{A C}$, the component of $\overline{A B}$ in the line of action of the force; because the motion from A to B can be accomplished by going from A to C and from C to B; and in this last portion no work is done. Since, then, when work is done, the motion must be in the direction of the force, if the work is done in producing change of motion of a body, the change produced must be in the *speed*, not in the direction of motion. Thus, when a body is moving in a horizontal circle with a

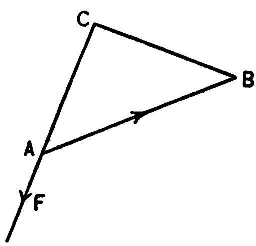


FIG. 54.

constant speed, no work is being done, and the energy remains constant.

The amount of work done depends upon force and distance in the direction of the force; consequently the numerical value of the work done may be defined as the product of the numerical values of the force and the distance in the direction of the force through which the motion takes place while the force is acting. The unit of work is, then, the product of one dyne and one centimetre, and is called the "erg." Since amount of energy is measured by the amount of work done or which can be done, the erg is also the unit of energy. Thus, if a body whose mass is m grams is raised vertically through a height h centimetres, the force overcome has been $m\,g$ dynes; and the work done is $m\,g\,h$ ergs. This work depends simply on the *vertical* distance h; being entirely independent of the path itself. For if there are two horizontal planes a distance h apart, the same amount of work is done in raising a given body from any point of the lower plane to any point of the upper along any path. Let $A\,B$ be a vertical line, $A'\,B'$ an inclined one, and $A''\,B''$ a curved one. The work done along $A'\,B'$ has just been proved to be the same as that done along $A\,B$; and since any curved line may be considered as the limiting form of a broken line consisting of a great number of extremely short straight lines, the work done along it is also the same. Since this work, $m\,g\,h$ ergs, has been done on the body, its energy has been increased by this much, owing to its change of position with reference to the earth. (If its final speed is different from its initial, there will be an additional change in its energy.)

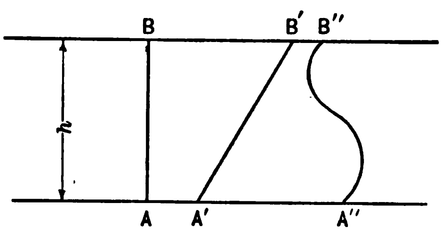


FIG. 55.

Again, consider a vertical spiral spring resting on a horizontal platform, and let it be compressed in some way, e. g. by means of one's hand. Work is done, because a force is overcome; and the spring gains as much energy as is lost by the body doing the work.

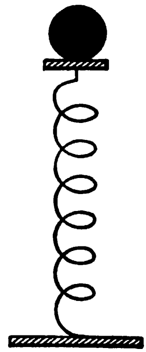


FIG. 56.

No idea can be formed at present of what the nature of the energy in these last two illustrations is; because it is not expressed in terms of matter and motion. (See Art. 7.) And such forms of energy as depend upon the positions of the bodies or upon the configuration of a body or system are called "potential energy." Thus any body has potential energy with reference to the earth; a compressed spring has potential energy; so has a twisted wire, a piece of iron near a magnet, a charged Leyden jar, etc. Ultimately, of course, it may be possible to describe any form of potential energy in terms of motion of matter, when the phenomenon comes to be better understood. So "potential energy" must be understood to be a term which is used to cover our present ignorance.

When work is done in changing the speed of a body, then, not alone is the amount of work known, but also the nature of the energy. Thus, let a body whose mass is m have its linear speed changed from s_0 to s, the direction remaining constant. By Formula (3), Article 18,

$$2\,a\,x = s^2 - s_0^2,$$

where a is the acceleration and x the distance required for the change. By definition, the force is $m\,a$; hence the work is

$$m\,a\,x = \tfrac{1}{2}\,m\,s^2 - \tfrac{1}{2}\,m\,s_0^2 \quad . \quad . \quad . \quad . \quad (30)$$

Consequently the energy has been increased by this amount. If the body had been at rest initially, i. e. if $s_0 = o$, it

would have had no energy of motion then; but, after the work $\frac{1}{2} m s^2$ was done, it would have had this amount of energy due to its linear speed, s. Therefore, when a body whose mass is m has a linear speed s, it has an amount of energy, $\frac{1}{2} m s^2$, owing to its motion. This energy which a body has in virtue of its being in motion is called "kinetic energy." A special illustration of kinetic energy is a body in rotation around a fixed axis. Consider the simplest case, that of a particle whose mass is m revolving around a point O in a circle of radius r with a constant angular speed ω. Its linear speed, s, $= r\,\omega$. Hence its kinetic energy is $\frac{1}{2}\ m r^2 \omega^2$. But $m r^2 = I$, the moment of inertia of m around O (Art. 43). Consequently, the kinetic energy is $\frac{1}{2} I \omega^2$. (This is perfectly analogous to the expression $\frac{1}{2}\ m s^2$ for translation.)

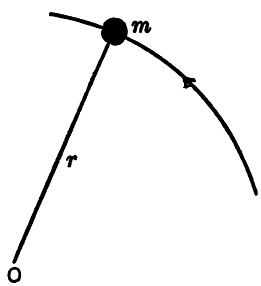


FIG. 57.

Let a force, F, act on m, perpendicularly to the radius r. The work done is $F x$. Let x be a *very small* distance in the direction of the motion of the particle; and let the corresponding angle be θ. Then $x = r\,\theta$; and the work $F x = F r\,\theta = L\,\theta$, where L is the moment around O. This again is perfectly analogous to $F x$ for translation.

69. Transfer of Energy. Work may be done by the passage of potential energy into potential, as when a compressed spring is made to raise a body from the earth; by kinetic energy passing into kinetic, as when a ball in motion strikes another ball and sets it in motion; by kinetic energy passing into potential, or *vice versa.* This last form of work is by far the most common. A body at a distance above the earth has potential energy; but if allowed to fall freely, it loses potential energy and gains kinetic; and of course the loss of one equals the gain of the other. Thus, if a body whose mass is m falls from any

point in one horizontal plane to any point in another horizontal plane at a vertical depth h below it, e.g. from B to A or B' to A' or B'' to A'', it will lose an amount of potential energy, mgh. Consequently, if the motion is along a smooth path, it will gain an equal amount of kinetic energy. That is, if s_0 is its speed at the upper plane, and s that at the lower,

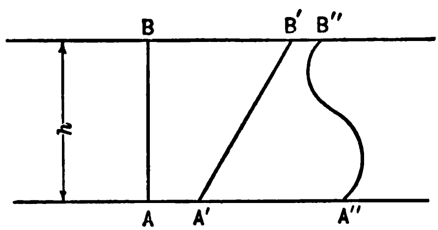


FIG. 58.

$$\tfrac{1}{2} m s^2 - \tfrac{1}{2} m s_0^2 = m g h,$$

or

$$s^2 - s_0^2 = 2 g h,$$

which is the formula previously proved in Article 18. This proves, then, that no matter how the body falls along a smooth path through a given vertical height, the change in the square of the speed is the same. A simple application of this principle is afforded when a particle falls along the inside of a vertical circle, or, what is the same thing, when a particle suspended by a cord from a fixed point is allowed to swing in a vertical plane. If the particle moves freely from P_1 downward along the circle, starting from rest, its speed at the bottom is such that

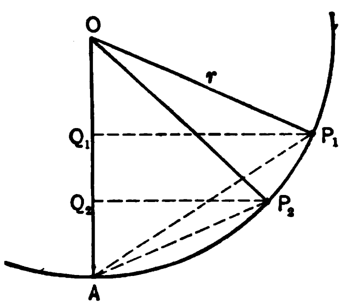


FIG. 59.

$$s_1^2 = 2 g \overline{Q_1 A}.$$

But by geometry it may be proved that

$$\overline{Q_1 A} = \overline{P_1 A}^2 / 2r,$$

where $\overline{P_1 A}$ is the *chord*, and r is the radius.

Hence
$$s_1 = \overline{P_1 A}\sqrt{\frac{g}{r}} . \quad . \quad . \quad . \quad . \quad . \quad (31)$$

If the particle had moved from P_2 down the circle, its speed at the bottom would have been

$$s_2 = \overline{P_2 A}\sqrt{\frac{g}{r}} .$$

This gives at once a method by which the speed may be measured in the experiment described in Article 29.

It is to be noted that a body left free to fall will always lose potential energy. Similarly a compressed spring free to move loses potential energy. A piece of iron separated from a fixed magnet has potential energy, but, if it is free to move, it loses it and gains kinetic energy. So in all cases. And it may be regarded as a general law that *there is a constant tendency in any system for its potential energy to decrease.* In fact, when in ordinary language it is said that there is a "force acting" between bodies, either an "attraction" or a "repulsion," all that is meant is that, if the bodies are left free to move, the motion will be such that the potential energy will decrease.

Work may be done at different rates, some bodies working rapidly, others slowly; and the rate of doing work — that is, the amount of work done in one second, if the rate is uniform — is called the "activity" or "power." The unit of activity is one erg per second, and 10^7 ergs per second is called a "watt." The standard sometimes used commercially is the "horse-power," which is the power required to raise 550 pounds through a vertical height

of one foot in one second. One pound is 453.6 grams, and one foot is 30.48 centimetres. Thus the force is $550 \times 453.6 \times g$ dynes; and the work done is $550 \times 453.6 \times g \times 30.48$ ergs. But this work is done in one second; hence

$$1 \text{ horse-power} = \frac{550 \times 453.6 \times g \times 30.48}{10^7} \text{ watts}$$
$$= 746 \text{ watts.}$$

70. Machines. Machines are mechanisms by means of which work is done, that is, energy transferred from one body to another. They themselves do no work; they neither gain nor lose energy, but simply transfer it. There are in most machines, however, parts which move over each other; and so some energy will always pass into these parts, producing "heat-effects" (see Art. 67). A perfect machine would be one which gave out again all the energy given it, without any "heat-effect." Of course the machine cannot deliver more energy than is furnished it; for this would be a violation of the principle of the conservation of energy. But the *force* which is doing the work need not equal the force produced; for let the work done on the machine be $F_1 x_1$, and that done by the machine be $F_2 x_2$. Then

$$F_1 x_1 = F_2 x_2,$$

i. e. $F_1/F_2 = x_2/x_1$; and so they may be different. This ratio F_2/F_1 of the force produced to the force working is called the "mechanical advantage" of the machine. The most common forms of machines are the lever, the pulley, and the screw.

71. The lever is simply a rigid beam pivoted at one point. Thus let a horizontal beam rest on an edge at O, and let a force, F_2, be applied at a point, A_2. If another force, F_1, is acting at another point, A_1, the problem is to

find what the numerical value of F_1 is when it can just overcome F_2 acting at A_2 (e. g. a man pulling down at A_1 wishes to raise a body hanging from A_2). Let the forces be parallel, and call the distances $\overline{OA_1} = l_1$, $\overline{OA_2} = l_2$. Then, if the beam is turned *slightly* around its pivot through an angle, θ, the point A_1 moves through a distance $x_1 = l_1 \theta$ in the direction of F_1; and A_2 moves a distance $x_2 = l_2 \theta$ in the direction opposite to F_2.

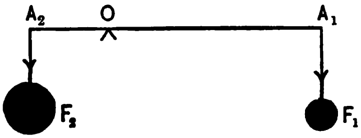


FIG. 60.

Hence, since
$$F_1 x_1 = F_2 x_2,$$
$$F_1 l_1 = F_2 l_2 \quad . \quad . \quad . \quad . \quad . \quad . \qquad (32)$$

That is, if F_1 and F_2 are in this ratio, they will just balance each other. If their ratio is different from this value, there will be acceleration.

Another way of solving the same problem is to consider it one of equilibrium, and deduce the condition that there shall be no moment. There are three forces acting: F_1, F_2, and the resistance of the pivot. In order to leave out this last, it is only necessary to take moments around the pivot; for then there are only two moments, $F_1 l_1$ and $-F_2 l_2$. Hence, if there is equilibrium, $F_1 l_1 - F_2 l_2 = 0$, or $F_1 l_1 = F_2 l_2$.

If the forces are not perpendicular to the lever, or if they are not parallel, the same demonstration holds if l is the perpendicular distance from the pivot to the line of action of the force.

Illustrations of levers are extremely common in machinery and in apparatus. One of its most important uses is in a so-called "chemical balance." In this apparatus a uniform horizontal beam is pivoted at its middle point, and carries suspended at its ends from two other pivots or

knife-edges two pans of equal masses. It is so arranged that its position of equilibrium is stable. If a body is now placed in each pan, there will be in general a move-

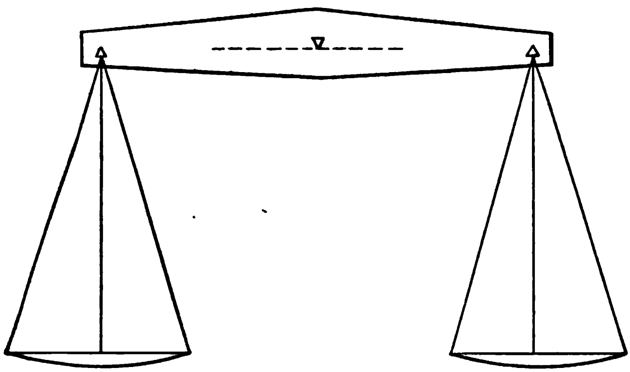


FIG. 61.

ment of the balance; but if the two forces produced by the earth on the two bodies are the same, the balance will not move; or, if displaced, it will make oscillations about its position of equilibrium. A balance can be used, then, to compare forces, and in particular to compare weights of bodies with reference to the earth.

72. A pulley is a grooved circular wheel free to turn around an axle passing through its centre. A cord is passed along the groove, so that a force applied to one end of the cord has the direction of its line of action changed. The amount of the force is not altered, if there is no friction; because, if F_1 and F_2 balance each other, their moments around the axis are $F_1 r$ and $-F_2 r$. Consequently, if there is equilibrium,

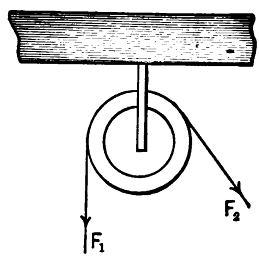


FIG. 62.

$$F_1 r - F_2 r = 0; \qquad \therefore\ F_1 = F_2.$$

Let the cord pass over a fixed pulley, then over a movable one, and back to a fixed support. Consider what force F_1 applied to the free end of the cord can balance the force F_2, acting downward on the movable pulley, when the two portions of the cord are parallel. Since a pulley simply changes the direction of a force, the force acting in each portion of the cord on the two sides of the movable pulley is F_1 acting up. Hence; if the movable pulley is in equilibrium,

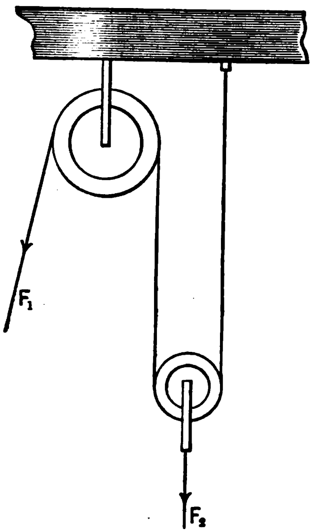


FIG. 68.

$$2F_1 - F_2 = 0, \text{ or } 2F_1 = F_2.$$

In actual practice, F_2 must be made to include the action of the earth on the pulley; and a correction must be made for all friction.

Another way of solving the question is to consider a small displacement such that the pulley is raised a distance x; that is, the string is shortened by an amount $2x$. Hence the work done by F_1 is $F_1 2x$, and that done on F_2 is $F_2 x$, and so

$$2F_1 x = F_2 x \quad \text{or} \quad 2F_1 = F_2.$$

If the portions of the cord are not parallel, the components of the forces in a vertical direction balance the force down. If acceleration is produced, then, of course, F_1 must be accordingly greater than $F_2 / 2$.

Pulleys are often made with more than one grooved wheel on the same axle. If the wheels can rotate independently of each other, they are called double, triple, etc. pulleys.

Consider an arrangement, as shown, where a cord passes over one wheel of a fixed double pulley, then over a movable single pulley, over the second wheel of the double pulley, and back to the axle of the movable pulley. There are thus three portions of the cord acting on the movable pulley, and they can be considered parallel. Consequently, as above, it is evident that

$$3 F_1 = F_2.$$

In general, then, if there are n portions of the same cord pulling up the lower pulley,

$$n F_1 = F_2 \quad . \quad (33)$$

Other combinations of pulleys than these with only one cord are possible; but, however many pulleys or separate cords are used, the mechanical advantage may always be calculated in a manner analogous to the one just given.

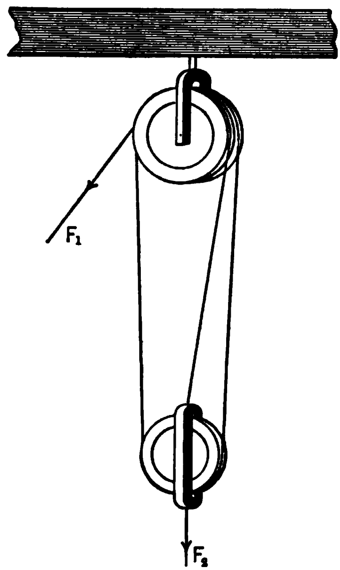


FIG. 64.

Some pulleys are made with two wheels on the same axle, but clamped rigidly together; the grooves are not smooth, but are cut into openings so as to receive the links of a chain which passes over them; and one wheel has, as a rule, one less opening than the other, and so has a smaller radius. Such a pulley is called a "differential" one. The chain is arranged as shown. It is generally continuous, and passes first over the larger wheel, then hangs in a free loop, then passes over the smaller wheel and back again to the larger wheel through a loop. A force, F_1, is applied to the chain just before it passes over the larger

wheel, and this is balanced by a force, F_2, applied in the loop formed just after the chain leaves the larger wheel.

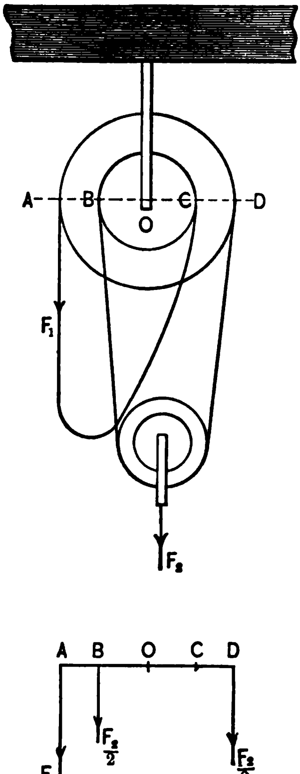


FIG. 65.

The chain cannot slip; and so, if the forces are parallel, they can be represented as shown in the diagram. Let the length of one link, that is, of one opening in the groove of the wheel, be a, and let there be n of them in the rim of the larger wheel and m in that of the smaller. Then the radius of the larger wheel is $na/2\pi$; that of the smaller, $ma/2\pi$. Hence, taking moments around the axle, since there is supposed to be equilibrium,

$$F_1 \frac{na}{2\pi} + \frac{F_2}{2}\frac{ma}{2\pi} - \frac{F_2}{2}\frac{na}{2\pi} = 0;$$

that is,

$$2F_1 n - F_2(n - m) = 0,$$

or
$$F_1 = F_2 \frac{n-m}{2n} \tag{34}$$

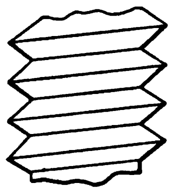
As stated above, in general m is made equal to $n - 1$; hence in this case

$$F_1 = F_2 \frac{1}{2n}.$$

So the mechanical advantage of a differential pulley is as great as desired, being conditioned only by the number of openings or teeth in the wheels.

73. A screw is simply a groove cut along the line of an inclined plane which is wound around a circular cylinder. The "pitch" of the screw is the angle between this plane

and a plane perpendicular to the axis of the cylinder, which may be called the base-plane; θ in the diagram is the pitch. A length l along the base-plane corresponds to a distance $l \tan \theta$ parallel to the axis. If the radius of the screw is r, one turn corresponds to a distance along the base plane, i. e. around the cylinder, of $2 \pi r$, and hence to a distance along the axis of $2 \pi r \tan \theta$. Let a moment, L, applied to the



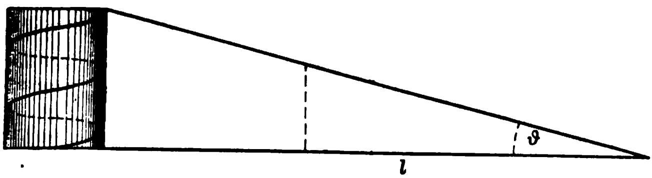


Fig. 66.

screw produce one complete turn, and let it just overcome a force, F, acting on a nut which is made to move along the screw. Then, the work done by the moment is L multiplied by the angle (Art. 57), which in this case is 2π; and that done against the force F is $F\, 2 \pi r \tan \theta$. Hence

$$L\, 2 \pi = F\, 2 \pi r \tan \theta.$$

And so

$$L = F r \tan \theta \quad . \quad . \quad . \quad . \quad . \quad . \quad (35)$$

Consequently, $1/\tan \theta$ is the mechanical advantage of a screw, and this can be as great a quantity as desired.

A differential screw can also be made by cutting threads of different pitches on the inside and outside of a hollow circular cylinder, and fitting a second screw inside this cylinder. Then, if the outer screw is turned through a

certain angle, the inner screw will move along the axis a distance which depends upon the difference of the two pitches. Its mechanical advantage is very great.

Comparison of Translation and Rotation.

Linear speed, $\frac{\text{distance}}{\text{time}}$;	angular speed, $\frac{\text{angle}}{\text{time}}$.
Linear velocity, linear speed in a particular direction;	angular velocity, angular speed around a particular axis.
Linear acceleration, $\frac{\text{linear velocity}}{\text{time}}$;	angular acceleration, $\frac{\text{angular velocity}}{\text{time}}$.
Mass;	moment of inertia.
Linear momentum, mass × linear velocity;	angular momentum, moment of inertia × angular velocity.
Force, mass × linear acceleration;	moment, moment of inertia × angular acceleration.
Linear impulse, force × time;	angular impulse, moment × time.
Work, force × distance;	work, moment × angle.
Kinetic energy, ½ mass × square of linear speed;	kinetic energy, ½ moment of inertia × square of angular speed.

WAVES

74. Wave-motion. By a wave is meant the advance of a disturbance into a medium; the portions of the medium do not move onward carrying the disturbance, but they hand it on from one portion to the next. Thus a stone dropped into a basin of water produces a wave, which spreads outward in all directions. If a stretched cord has one end moved suddenly sideways, a wave will pass along the cord. If a stretched spiral spring has one end moved suddenly in the direction of the length of the spring, a

wave of compression or rarefaction will move along the spring. In all these cases a certain state of affairs, a certain condition, advances and forms the wave. The individual particles of the medium move only a short distance. If stones are dropped into the water at regular intervals, or if the end of the string is moved sideways and back regularly, or if the end of the spiral spring is moved forward and back regularly, there will be a succession of waves. The waves on the surface of the water will be considered more in detail. They consist of a

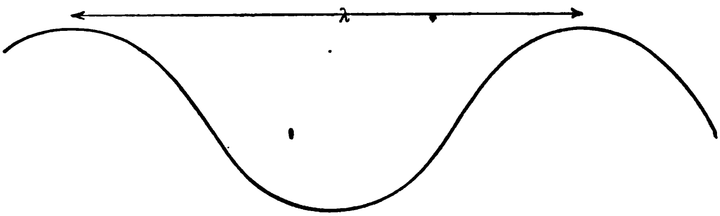


FIG. 67.

series of crests and hollows, having at any instant, in the simplest case, the appearance of a sinuous curve. The advance of the wave consists in the moving forward of this form. The distance from crest to crest is called the "wave-length;" the number of crests which pass any point in one second is called the "wave-number" or the "frequency;" the distance through which the wave advances in one second is called its velocity. The symbols for wave-length, wave-number, and velocity are λ, n, u; and consequently

$$u = n\lambda \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (36)$$

The individual particles of the water are moving in vertical planes, making harmonic vibrations. Adjacent particles are of course in different positions of their motions; but there is always at a distance of one wave-length from any particle another particle whose motion is identical with its

own, for at an interval of a wave-length everything is repeated.

Similarly, as a wave advances along a string, there are crests and hollows; and the definitions of wave-length, wave-number, and velocity are the same as above. The individual particles of the string move across the length of the string, making harmonic vibrations. Such a wave as this is called a "transverse" wave, because the motion of the particles is perpendicular to the direction of advance of the wave.

In the spiral spring there are condensations and rarefactions at regular distances apart; and the wave-length is the distance from one condensation to the next, etc. In this wave the individual portions of the spring are making harmonic vibrations backward and forward in the line of the advance of the wave. Such waves are called "longitudinal."

If two or more waves are passing through a medium at the same time, the resultant effect at any point of the medium is simply the geometrical sum of the separate effects which each wave would by itself have produced, provided only that the motion of the particles of the medium is small compared with the length of the waves. Any wave, however complicated, may be considered as composed of a number of the simplest waves, whose form is that described above.

It requires energy to produce waves, and, if a body stops waves, it receives energy. So waves can be regarded as carrying away energy from a source and giving it to whatever destroys the motion. The energy in a wave is, naturally, partly kinetic and partly potential; because the portions of the medium are in motion, and they are also displaced with reference to each other.

If the wave-length is long, the waves will pass around a small obstacle, as water-waves do around a small pole sticking up out of the surface; whereas a large obstacle

will either stop them or reflect them. If an object stops the waves, it is said to "absorb" them or to absorb their energy.

The amount of energy carried by the waves through an area of one square centimetre placed at any point perpendicular to the direction of the waves is called the "intensity" of the waves at that point. If the waves are sent out in an isotropic medium from a point, they will spread outward in the form of spheres. If two spheres of radii, r_1 and r_2, be drawn around the point-source of the waves, the same amount of energy must be passing through these two spherical surfaces at any instant, as it is carried outward by the waves. The areas of the surfaces of the two spheres are $4\pi r_1^2$ and $4\pi r_2^2$. So, if E is the total amount of energy passing through each of the surfaces, I_1, the intensity at the distance r_2 is $\frac{E}{4\pi r_1^2}$; and I_2, the intensity at the distance r_1, is $\frac{E}{4\pi r_2^2}$. Hence

$$I_1 : I_2 = \frac{1}{r_1^2} : \frac{1}{r_2^2} \quad \ldots\ldots \quad (37)$$

That is, the intensity of waves in an isotropic medium varies inversely as the square of the distance from the source.

It may also be proved that the intensity of the waves varies directly as the square of the amplitude of the vibrations of the particles of the medium which carries the waves.

Since liquids and gases take the shape of the vessels which contain them, they do not resist any attempt made to displace one layer over another; and so a wave like a transverse wave in a cord is impossible in them. The ordinary waves on the surface of water or of any liquid are not due to any reaction of the liquid itself, but are caused by the effect of the earth's force of gravitation

(Art. 76). Liquids and gases, then, of themselves can transmit only longitudinal waves; while solids can transmit both longitudinal and transverse ones.

The velocity of any train of waves in any homogeneous medium must vary directly as the elasticity of the medium and inversely as its inertia. It may be proved mathematically that this is the case, and that the velocity depends upon nothing else, if there are no obstacles immersed in the medium, which may affect waves differently. The velocity is the same for waves of all lengths, short and long; and of all amplitudes, provided only that the limits of elasticity are not reached. These facts are not true of waves on the surface of water, because they are not due to any elastic properties of the water; and in fact long water-waves travel faster than shorter ones. Further, if the medium is changed in any way, as by immersing a great number of small bodies in it which affect waves differently according to their wave-numbers, the velocity of different waves is no longer the same.

75. If a wave of any kind is sent out in any medium from a centre of disturbance, its progress is marked by the motion of the portions of the medium when the wave reaches them; and the locus of the points which the wave has just reached is called the "wave-front." In the case of a wave proceeding out from a point into an isotropic medium, the wave-front is obviously a sphere, and the wave is called a "spherical" one. If the centre is very far away, the sphere is so large that any small portion of the surface is practically a plane; and so we can speak of a "plane" wave.

Consider a spherical wave whose centre of disturbance is O. At a certain instant its front passes through a series of points, P_1, P_2, P_3, etc., on a sphere of radius R; and they are set in vibration. At an interval of time later the wave-front reaches another series of points, Q_1, Q_2, Q_3, etc., on a sphere of radius S. It is obvious from ordi-

nary geometry that, if spheres of the same radii, $S-R$, are described around each of the points P_1, P_2, P_3, etc., they will just touch the outer spherical surface on which lie the points Q_1, Q_2, Q_3, etc. Consequently, the effect is just the same at the outer spherical surface if the spherical wave itself advances there, or if the original wave be considered *replaced* at any instant by a great number of smaller spherical waves sent out by each of the points P_1, P_2, P_3, etc., of the original wave-front at that instant; because the time taken for the original wave to pass from the first position to the second is

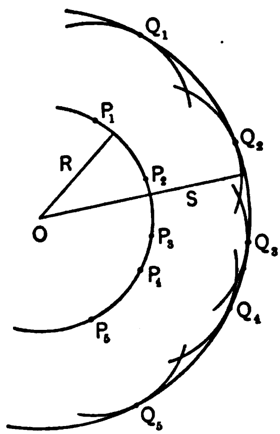


FIG. 71.

$\dfrac{S-R}{u}$, where u is the velocity of the wave; and in this same time the smaller spherical waves from P_1, P_2, P_3, etc., will just reach the outer surface. This process is said to be "breaking up" a "primary" wave into "secondary" waves. Although the effect at the wave-front is the same, however it is described, in terms of primary or secondary waves, it is often convenient from a mathematical standpoint to refer the disturbance to the secondary waves rather than to the primary one.

CHAPTER III

GRAVITATION

76. Weight. Reference has been made several times to the fact that all bodies have their motion influenced by the presence of the earth; and it was stated that any body falling toward the earth has a constant acceleration which is the same for all bodies at any one place on the earth. This constant acceleration was called g; and its numerical value is nearly 980 in the C. G. S. system. This fact, that all bodies, whatever their material or mass, fall towards the earth with the same acceleration, is most remarkable. It may, however, be proved by experiment in many ways. 1. Allow bodies of various kinds and masses to fall inside a vacuum. All resistance of the air is now removed, and the bodies are observed to fall side by side. 2. It was proved in Article 51 that the period of a pendulum was $T = 2\pi\sqrt{\frac{l}{g}}$; and so $g = \frac{4\pi^2 l}{T^2}$; l and T can both be accurately measured; and thus g may be determined. It is found that the same value of g is thus obtained when the pendulum is made of any body or has any mass. Consequently the force "acting on" any body of mass, m, towards the earth is mg; and this force is often called the "weight" of the body with reference to the earth.

Since a chemical balance can measure forces, it can compare the weights of two bodies. Place one body in each pan of the balance, and alter the mass of one by adding to it or taking away, until the two weights are equal. Let

the masses, as measured by inertia, be m and m_1. Then, as their weights are equal,

$$m g = m_1 g,$$

and consequently $m = m_1$, since at the same place g is the same for all bodies; and so the masses as measured by inertia are also equal. This is equivalent to saying that, if two bodies have the same mass as measured by inertia, they also have the same weight. Consequently equal masses might perfectly well be defined as corresponding to equal weights instead of as corresponding to equal inertias: the two definitions are identical since g is proved by experiment to be a constant, the same for all portions of matter.

For different points of the earth's surface g is different; and so the weight of a body varies, although its mass does not. But, if two bodies have the same weight at any one place, they will also agree at any other place. The chief causes of the variations in g are the rotation of the earth, the ellipsoidal shape of the earth, and local disturbances such as mountains and differences in height above sea-level.

77. Universal Gravitation. This effect of the earth on a body near it is only a special case of a more general law of nature. There seems to be undoubted evidence that each portion of matter in the universe influences the motion of every other portion; and a law has been proposed which seems to be in accord with all the observed facts, except possibly when the portions of matter are as small and as near together as are the molecules of a body. This law, known as Newton's Law of Universal Gravitation is, that, if two bodies whose masses are m_1 and m_2 are at a distance, r, apart, where r is large compared with the size of the bodies, their mutual action is such that, if free to move, they approach each other with a force which

is proportional to the product of their masses and inversely to the square of their distance apart, i. e.

$$F = \gamma \frac{m_1 m_2}{r^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where γ is simply a constant of proportionality.

It may be proved without much difficulty that, if either of the bodies is a homogeneous sphere, its action is exactly what would be produced if the entire mass was concentrated at its centre.

Ordinary weight with reference to the earth is an illustration of the law; because, if m_1 is the mass of the body, m_2 that of the earth, and r the distance from m_1 to the centre of the earth, $\gamma m_2 / r_2$ is a constant entirely independent of m_1. Hence $F = m_1 g$. Further, the forces due to the earth "acting on" different portions of matter near each other are all parallel, because the centre of the earth is so far away from the surface that lines drawn to it from bodies near each other on the surface are parallel.

The law has been directly verified in many ways: (1) by direct measurement of the forces between different masses at varying distances; (2) by the application of deductions from it to astronomical calculations and predictions. An application of this last method is to predict the period of the revolution of the moon around the earth. In Article 19 it was proved (Formula 8) that the period of a body moving in a circle of radius r, when the acceleration towards the centre is a, is

$$T = 2\pi\sqrt{r/a}.$$

Call the mass of the earth m_2, that of the moon m_1. Then the acceleration at the moon due to the earth is

$$a = \frac{F}{m_1} = \gamma \frac{m_2}{r^2},$$

where r is the radius of the moon's orbit.

Similarly, the acceleration at the surface of the earth is $\frac{\gamma m_2}{r_1^2}$, where r_1 is the radius of the earth. But r, the distance from the centre of the earth to the moon, is very nearly 60 r_1; and the acceleration at the surface of the earth is known to be nearly 980. Hence

$$980 = \gamma \frac{m_2}{r_1^2}; \quad a = \gamma \frac{m_2}{60^2 r_1^2}.$$

Eliminating γ,

$$a = \frac{980}{3600} = 0.272.$$

Therefore, $$T = 2\pi \sqrt{\frac{r}{0.272}}$$

$$r = 60\, r_1 = \frac{60 \times 4 \times 10^9}{2\pi},$$

and so $T = 27$ days 6 hours, whereas the observed period is 27 days 8 hours, which is very close agreement.

The formation of the tides in the oceans is also an illustration of the influence of matter upon matter; in this case, of the sun and moon upon the fluid portions of the earth; and the main features of the tides may be easily predicted by elementary applications of the general law of gravitation. There are, of course, innumerable other illustrations of the truth of the law.

78. Centre of Gravity. There is, then, a force due to the earth "acting on" every portion of matter; and, if the matter is near the earth, this force has for its value $m g$, and the forces acting on the different portions of matter of the same system are parallel. As was proved in Article 55, if parallel forces act on a system of bodies, each force having for its value $m g$, the resultant force has for its value $M g$ where M is the sum of the separate masses, and its line of action is parallel to the component forces and passes through the centre of inertia. A special case of a

system of particles is a single large body; and so in all cases there is a certain point, viz. the centre of inertia, at which the whole resultant force may be supposed to act. For this reason the centre of inertia of any body or system is often called its "centre of gravity;" and its position may be calculated as was shown in Article 40.

In any actual case it may be found by experiment. Suspend the body so that it is free to turn around a horizontal axis at any point, P; and let C be the centre of inertia. There is a force, mg, vertically downward through C; and this force will have a moment around the axis at P unless C is in a vertical line through P. Consequently, if C is not in that line, the body will turn around the axis until it does come there. (In general the body will make oscillations about the axis, and such a vibrating body is called a "compound" pendulum, to distinguish it from a simple pendulum. See Article 51. If there is friction, the body will soon come to rest.) So it is known that, if the body is at rest, C must be somewhere in a vertical line passing through P. Now suspend the body so that it is free to turn around another horizontal axis not in this line; and, when it comes to rest, the centre of inertia must be somewhere in a vertical line passing through the new point of support. Since, then, the centre of inertia must lie on both these lines, it must be the point of their intersection.

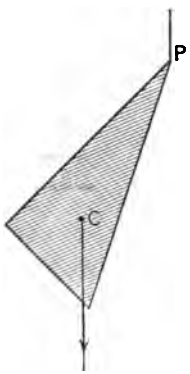


FIG. 69.

If a body is so balanced that when it is at rest the centre of gravity lies vertically *above* the axis, it is in unstable equilibrium; because, if it is displaced at all, it will immediately turn over so that its centre of gravity may get as near the earth as possible. It will come to rest again (if there is friction) with its centre vertically

below the axis, where the equilibrium is stable. This is an illustration of the general principle that the potential energy of a system always tends to become as small as possible. If the axis passes through the centre of inertia, the body is in neutral equilibrium; for any displacement produces a constant permanent effect. In general, therefore, a body is in stable equilibrium with reference to the earth, if its centre of gravity is as near the earth as it can get under the existing conditions of restraint.

Units

Length: centimetre (cm.).
  1 cm. = 0.39370 inches = 0.032809 feet.
  1 in. = 2.5400 cm.

Mass: gram (g.).
  1 g. = 0.0022046 pounds.
  1 lb. = 453.59 g.

Time: mean solar second (sec.).

Speed: 1 cm. per sec.

Acceleration: (1 cm. per sec.) per sec.

Force: dyne. 1 g. given unit acceleration in 1 sec.
  Mass m. has weight mg. dynes.
  1 gram weighs g. dynes; 1 pound weighs 4.45×10^5 dynes.

Energy: erg. 1 dyne × 1 centimetre.
  1 foot-pound = 1.3562×10^7 ergs.

Activity: Watt, 10^7 ergs. per sec.
  1 horse-power = 7.46×10^2 Watts.

Acceleration due to the earth.

Boston,	980.4.	Denver,	979.6.
Philadelphia,	980.2.	San Francisco,	979.95.
Washington,	980.1.	Greenwich,	981.17.
Chicago,	980.3.	Paris,	980.96.
St. Louis,	979.99.	Berlin,	981.24.

CHAPTER IV

PROPERTIES OF SIZE AND SHAPE OF MATTER

79. Introduction. Both these general properties of matter, inertia and weight, are common to all forms of matter, solids, liquids, and gases; and all the laws which have been stated apply equally well to them all. There are, however, certain other properties which are different for various forms of matter. As noted in Article 3, the name "solid" is commonly given to a body which has a volume and a shape of its own; the name "liquid," to a body which has a volume of its own, but which assumes the shape of the containing vessel; the name "gas," to a body which has neither volume nor shape of its own, but assumes those of the containing vessel. These definitions are by no means perfectly exact; but they satisfy all general purposes.

The volume and shape of any body may be changed; and the behavior of the body after these changes are produced distinguishes definitely between different forms of matter. If either the shape or the volume of a solid is altered, there is in general a tendency for the body to return to its previous condition; and so work is required to produce the change, and, if the cause producing the change ceases to act, the body will return to its previous state. Various solids differ widely as to these properties. If the volume is changed with ease, the solid is said to be "compressible;" while, if the shape is changed with difficulty, the body is said to be "rigid." Thus a piece of steel is not easily compressible and is rigid; a piece of

cork is very rigid; a piece of rubber is compressible and not rigid. In all these cases, though, the body, if not changed too much, will return to its previous state if allowed to; and such a body is said to be "elastic." In some solids, however, when a change is produced, it becomes partially or completely permanent; thus, after a piece of lead or putty is changed in shape or volume, it does not completely recover its former condition. The molecules of the body have been permanently displaced over each other; and part of the work done has been produced by the passage of energy into the molecules, and so a "heat-effect" is produced. Such bodies are called "plastic" or "inelastic."

A perfect liquid would offer no resistance at all to a change in shape; that is, one layer could move over another with perfect freedom. But no actual liquid has this property; and so, when one portion of a liquid moves over another there is always a resistance which tends to make the motion of the two portions the same. This property is called "viscosity." Any force, though, however small, will make one portion of the liquid move over another, if only it is applied long enough. All liquids are more or less compressible; and all tend to recover their former volume when the force is removed; that is, all liquids are elastic.

A perfect gas would also not be viscous; but all actual gases are more or less so. All gases are compressible; and they offer a resistance to a decrease in volume. If the volume is restored to its original value, the gas again expands; and so it may be said to be elastic.

Some solids and all liquids and gases are, then, elastic, if the change produced is not too great; and this limiting change beyond which a body cannot go without losing the power to recover its previous condition is called the "limit of elasticity." The change is always produced by some force; and the amount of the change depends upon the

amount of the force, the area over which it acts, the amount of the property which is changed, and the nature of the body itself. To simplify this statement, various definitions have been adopted. The internal "stress" is the force of reaction produced in the body by the applied force, divided by the area over which the reaction acts. The "strain" is the amount of the change, divided by the amount of the quantity changed. The "coefficient of elasticity" for any particular kind of change in any body is the ratio of the stress to the strain. And experiments prove that for all strains within the limits of elasticity, the coefficient of elasticity corresponding to any particular type of change is a constant for any one substance. This is called "Hooke's Law." It is equivalent to saying that the strain is proportional to the stress.

Special Properties of Solids

80. A solid has both a volume and a shape of its own; and therefore any elastic solid has two independent coefficients of elasticity, — one corresponding to a change in volume, the other to a change in shape.

81. Change in Volume. To change the volume of a solid without changing its shape or those of its smallest portions is not easy; for it is necessary to apply a uniform stress all over the body perpendicular to its surface at each point. It can be done, however, by immersing the solid in a liquid and then increasing the pressure of the liquid. In defining the coefficient of elasticity it was necessary to specify the *internal* stress, because inelastic bodies have no internal reaction when a force is applied to change them. But, if a body is elastic, there is a reaction; and the internal force equals the external force when a change is produced. Let the stress producing the change in volume be p where p is the force divided by the area; and let the original volume be v, and the change pro-

duced by p be Δv. The strain is then $\Delta v / v$, and the coefficient of elasticity for a change in volume is, therefore,

$$k = \frac{p}{\Delta v / v} = \frac{v p}{\Delta v} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

This coefficient is sometimes called the "bulk-modulus;" and it has, of course, different values for different solids. Some of the values are, —

For steel, $k = 18.4 \times 10^{11}$;
" flint glass, $k = 4.15 \times 10^{11}$;
" cast-iron, $k = 9.64 \times 10^{11}$.

82. Change in Shape. To change the shape of a solid without altering its volume sensibly is not difficult. Nail firmly two boards to opposite ends of a block of wood; then, holding one board firmly fixed, push the other sideways in the direction of its own plane. The shape will be changed, as shown; the volume, however, will remain sensibly the same. If the force is F, and the area of the top of the block A, the stress is F / A. Call this T. A measure of the strain is the angle, θ, between the former direction of the edge of the block and the new one produced by the stress. The coefficient of elasticity for a change in shape is, then,

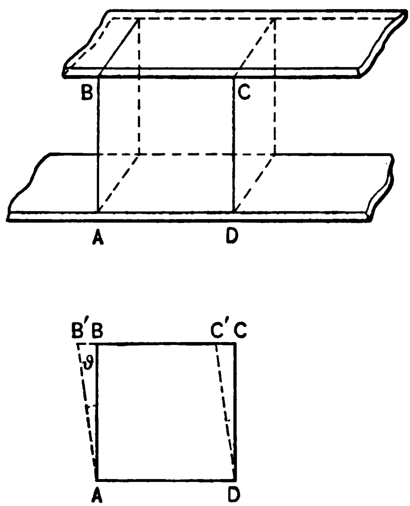


FIG. 70.

$$n = T / \theta \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

This coefficient is sometimes called the "coefficient of rigidity."

A change in shape is also produced, without a change in volume, when a rod or wire is *twisted* around its axis of figure. This is easily seen if the alteration in any small cubical portion of the rod or wire is considered. If one end of the rod or wire is held firm, and the other twisted by a moment L around the axis of figure, the angle θ through which the lower end will be turned is given by the following formula, which may be deduced by the help of higher mathematics: —

$$\theta = \frac{l\,L}{n\,B^4}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where l is the length of the rod or wire and B is a constant for any one rod or wire, depending upon its dimensions. For a circular cylinder of radius r, $B^4 = \pi\,r^4\,/\,2$; and so, for a wire of circular cross-section,

$$\theta = \frac{2\,l\,L}{\pi\,n\,r^4} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

If such a wire be twisted and then set free and allowed to make torsional vibrations, the moment tending to *oppose* the motion at any instant, due to the reaction of the wires will be

$$L = \frac{\pi\,n\,r^4}{2\,l}\,\theta.$$

Consequently, if a disc whose moment of inertia around the axis is I is fastened to the free end of the wire, its angular acceleration at any instant of the vibration will be $L\,/\,I$ or $\frac{\pi\,n\,r^4}{2\,l\,I}\theta$; and so the vibrations will be harmonic (see Arts. 25 and 51), and the period of one complete vibration will be

$$T = 2\,\pi\sqrt{\frac{2\,l\,I}{\pi\,n\,r^4}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

The quantities T, l, I, r, can be easily measured; and so n, the coefficient of rigidity, may be determined for the substance of which the wire is made. It has been found that

For steel, $n = 8.2 \times 10^{11}$;
" glass, $n = 2.4 \times 10^{11}$;
" cast-iron, $n = 5.3 \times 10^{11}$.

The ordinary vibrations of a spiral spring are due to the fact that, when the spring is elongated, the wire itself is *twisted*; and so the vibrations are really torsional ones.

83. In general, unless special precautions are taken, there will be a change in both volume and shape; and one of the most common alterations is that when a rod or wire is compressed or stretched by forces applied, parallel to the axis of figure, at the two ends. By considering any small cubical portion of the rod or wire, it will be obvious that its shape is changed; and the volume will also be in general. Thus, let a wire having a circular cross-section of radius r, and a length l, be fastened firmly at one end and be stretched by a force F applied at the lower end. Let the change in length produced by F be Δl. Then the stress $p = F / \pi r^2$; the strain is $\Delta l / l$; and the coefficient of elasticity, which in this case is called "Young's modulus," is

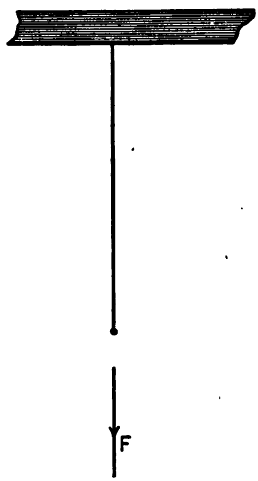


FIG. 71.

$$E = \frac{p}{\frac{\Delta l}{l}} = \frac{Fl}{\pi r^2 . \Delta l} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

These quantities can all be easily measured; and so E can be determined.

For steel, $E = 21. \times 10^{11}$;
" glass, $E = 6. \times 10^{11}$;
" cast-iron, $E = 13.5 \times 10^{11}$.

Since E is a constant for any one substance for slight changes, Δl is proportional to F; that is, the change in length is proportional to the stretching force. The same statements and formula apply also to the case of a rod being compressed by forces applied at the two ends, e. g. a pillar supporting a building.

Since, in determining Young's modulus, there is a change in both shape and volume, E must depend upon k and n; and it may be proved from theoretical considerations that

$$\frac{1}{E} = \frac{1}{9k} + \frac{1}{3n} \quad \cdots\cdots \quad (7)$$

Consequently, if E and n are determined for any one substance, k may be calculated; and in fact this is the way its value is nearly always found. (The nature of a sub-

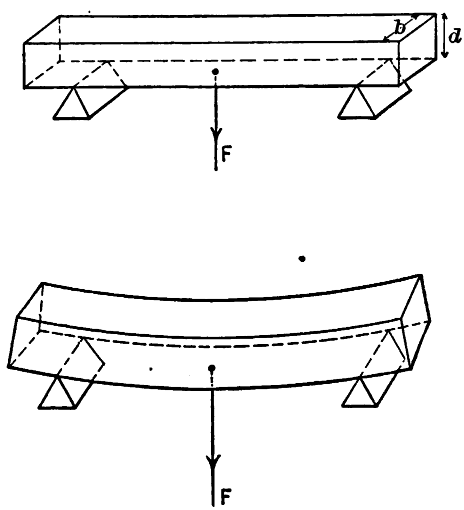


FIG. 72.

stance is often changed when it is drawn out into a wire; so the calculated value of k applies to the substance in the form of a wire.) Another illustration of a change in both volume and shape is when a rod is bent. If a rod of breadth b and depth d is placed upon two knife edges at a distance l apart, and if a force, F, parallel to the depth of the rod, is applied at a point half-way between the knife edges, the displacement of that point in the direction of the force may be proved to be

$$\Delta = \frac{F l^3}{b\, d^3} \frac{1}{E} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (8)$$

E is Young's modulus; and so it can be determined by this experiment as well as by the one described above.

Special Properties of Liquids

84. Introduction. The distinguishing characteristics of a liquid are: 1. Any force, no matter how small, can cause one portion to move over another. 2. It is only slightly compressible, and so keeps a definite volume. As a consequence, a liquid may be considered as occupying a certain volume and so having a free surface. 3. Its smallest portions or molecules are in motion in all directions within the volume, on the kinetic theory of matter.

These ideas of the properties of a liquid are derived, of course, from observation and experiment; and every deduction from these properties is found to be true. A few of the most important of these deductions are the following.

85. Properties of a Liquid at Rest. 1. The thrust or force against any surface in contact with a liquid which is at rest (that is, not flowing) is always perpendicular to it. For, if the thrust were not perpendicular, it would have a component parallel to the surface; and this component would produce flowing of the liquid, which is contrary to the hypothesis that it is at rest. (See Fig. 73.)

86. 2. At any point in a liquid at rest the pressure is the same in all directions. By "pressure" is meant the force acting over a surface divided by the area of the surface; i. e. if it is uniform, it is the force per square centimetre. So the pressure in any direction at a point is the force acting on an infinitesimal surface at the point, perpendicular to that direction, divided by the area of the surface as it is made smaller and smaller.

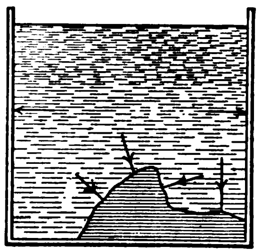


FIG. 73.

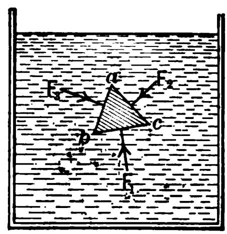


FIG. 74.

Consider the forces acting on any small portion of the liquid enclosed in a volume whose cross-section is a triangle $\Delta\, a\, b\, c$, and whose height is any small distance. Call the three forces perpendicular to the three faces F_1, F_2, F_3, and the areas of the faces A_1, A_2, A_3. Then the three pressures, p_1, p_2, p_3, are F_1 / A_1, F_2 / A_2, F_3 / A_3. If the volume is made small enough, these three forces may be considered as acting at a point; and, since by hypothesis the liquid is not flowing, the sum of the components of the forces in any direction must equal zero. Take this direction parallel to the face on which F_3 is acting. The component of F_1 is $F_1 \sin (a\,b\,c)$; that of F_2 is $-F_2 \sin (b\,a\,c)$; that of F_3 is 0.

Hence $$F_1 \sin (a\,b\,c) - F_2 \sin (b\,a\,c) = 0.$$

Hence, substituting the values of the two corresponding pressures,

$$p_1 A_1 \sin (a\,b\,c) = p_2 A_2 \sin (b\,a\,c).$$

But by geometry, $A_1 \sin(a\,b\,c) = A_2 \sin(b\,a\,c)$,
and so $p_1 = p_2$.

Since the triangle taken is any triangle, the two pressures, p_1 and p_2, are pressures in any two directions; consequently the theorem is deduced. (This theorem may also be proved to be true of a liquid which is flowing.)

87. 3. The free surface of a liquid at rest is perpendicular to the forces acting. Thus any surface of liquid, not too small, on the earth, is perpendicular to the force of gravity; for, if it were not so, the force would have a component parallel to the surface, and there would be a flow of the liquid. (Where a liquid meets a solid, the surface is not horizontal but curved, for reasons to be explained later.)

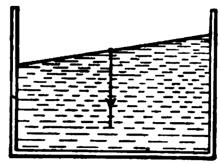


FIG. 75.

88. 4. The pressure at any point in a liquid at rest is due to two causes: (*a*) the containing vessel, (*b*) the fact that layers near the bottom of a liquid must support the weight of the liquid above them.

Consider these two causes separately. To do away with the second, the experiments may be considered as being performed at the centre of the earth. If the liquid is enclosed in a vessel which it completely fills, the pressure must be the same at all points throughout the liquid as at all points against the walls of the vessel. This is a direct consequence of the fact that the molecules of the liquid are moving freely about in all directions; and so, if the pressure were greater at one point than another, the liquid would immediately flow so as to make the pressures the same.

An illustration of this fact is the "hydraulic press," which in principle consists of two cylinders of different radii, closed by pistons and connected by a tube. Let the entire space be filled with a liquid, and call the areas of

the two pistons, A_1 and A_2. If a force, F_1, acts on A_1, the force F_2 which must be applied to A_2 in order to balance F_1 is

$$F_2 / A_2 = F_1 / A_1,$$

since the pressures must be the same. Consequently a small force acting over a small area may produce as great a force as is desired simply by increasing the area on which it acts. This is not contrary to the conservation of energy, because, if the smaller force produces any motion, the larger force will be overcome through a distance smaller in proportion.

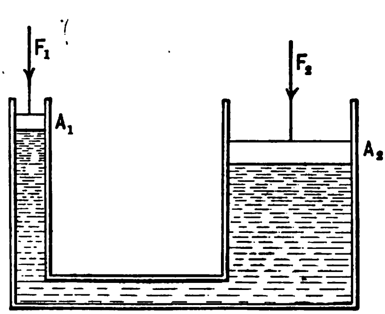


FIG. 76.

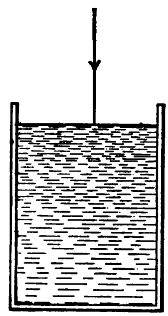


FIG. 77.

In the case of a vessel of water with a free surface, such as is ordinarily dealt with, the weight of the air on the surface produces a uniform pressure throughout the liquid just as if the vessel was closed by a piston which was pressed down by a force equal to the weight of the atmosphere on the surface. (If the surface is very small, there is an additional pressure called the capillary pressure, which will be discussed later on. See Article 96.)

89. Pressure due to Weight. In addition to this pressure due to the surface, there is the additional one, at any point in a liquid, due to the weight of the liquid above it. If the liquid is not flowing, consider the forces acting on any surface whose area is A, which is parallel to the free sur-

face but at a depth h below it. The force down on this surface is equal to the weight of the atmosphere on an area A of the upper surface, vertically over the lower one, plus the weight of the liquid above it. The volume of this vertical column of liquid is $h\,A$; hence, if ρ is the density of the liquid, the mass is $\rho\,h\,A$, because by definition the density is the mass divided by the volume; and so the weight is $\rho\,g\,h\,A$. Consequently, the force downward, due to the weight of the liquid, is $\rho\,g\,h\,A$; and the pressure downward therefore is $\rho\,g\,h$. But, as stated above, the pressure at any point of the liquid is the same in all directions. Hence the pressure at any point, due to the liquid itself, is

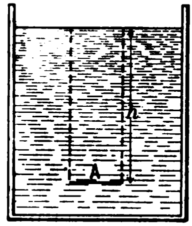


FIG. 78.

$$p = \rho g h \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

The entire pressure at any point at a depth h below the surface is, then, the sum of the pressure on the surface and $\rho\,g\,h$.

It is seen that this last pressure, $\rho\,g\,h$, does not depend in the least on the shape of the containing vessel, but only on the depth below the free surface.

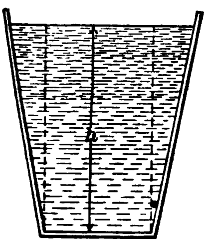


FIG. 79.

If a vessel having the shape of a frustum of a cone with its smaller end closed, is filled with a liquid to the height h, the pressure on the bottom, due to the weight of the liquid, is $\rho\,g\,h$. Hence, if the area of the bottom is A, the entire force on the bottom, due to the liquid itself, is $A\,\rho\,g\,h$, which is the weight of a cylindrical column of the liquid of cross-section A and height h. The weight of the rest of the liquid is borne by the walls.

If the vessel which contains the liquid has the shape of a frustum of a cone with its larger end closed, the pressure on the bottom, due to the liquid, is $\rho g h$; and if the area of the bottom is A, the entire force produced on it is therefore $A \rho g h$. The weight of the entire liquid is not this much; but the extra force comes from the reaction of the walls downward. The liquid exerts a thrust perpendicular to the walls, consequently there is an equal reaction against the liquid; and in this case there is a component vertically downward.

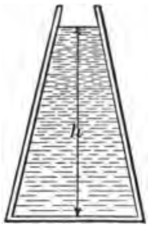


FIG. 80.

90. If, then, the same liquid is standing in a series of connecting tubes of any shape, the level must be the same in them all. Let the base of the tubes be horizontal. The pressure must be the same at all points in the liquid along this base; for, if it was not, there would be a flow. But the pressure below each tube, due to the liquid itself, is $\rho g h$; and as ρ and g are the same for all the tubes, h must be also. That is, the free surfaces must be at the same height above the horizontal base; and so they are all in the same level. Also, taking a series of points in each tube at the same depth below the surfaces, they will all be in the same level; and the pressures are the same for all the points. Hence, in connecting tubes containing the same liquid, the pressures at all points in the same level are the same.

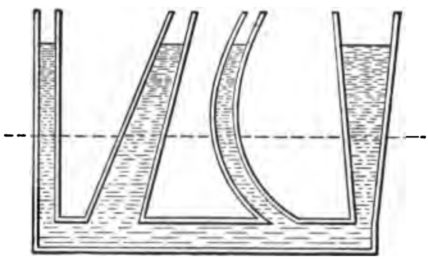


FIG. 81.

Another effect of the influence of weight upon liquids is that, if a number of liquids which do not mix or react on each other, and which are of different densities, such as

mercury, oil and water, are placed in a vessel, they will arrange themselves so that the surfaces between them are horizontal, and in such an order that the densities decrease from layer to layer as one goes from the bottom up.

If a liquid is contained in a vessel of any shape, there is a thrust against the walls due to the pressures at each point. The pressure at the free surface, due to the liquid itself, is zero; that at any other point at a depth h' is $\rho g h'$; so that the pressure increases uniformly, and the *average* pressure in a liquid of depth h is $\frac{1}{2} \rho g h$. The thrust or force against the wall is the resultant of all the individual forces acting on each point of the wall; and the point of application of this resultant is called the "centre of pressure." Its position can generally be determined by calculation, if the shape of the wall is known. If the wall is straight and has the same width from top to bottom, this resultant thrust is the average pressure multiplied by the area of the wall.

91. Measurement of Density of a Liquid. If different liquids which do not mix or react on each other are placed in connecting tubes, their free surfaces will not stand at the same height. Let two such liquids of density, ρ and ρ_1, be placed in the two arms of a U-tube. The denser liquid will occupy the lower connecting tube, and, if there is enough of it, will rise to a certain height in one arm, while the lighter liquid will be in the other arm. Imagine a horizontal plane passed through

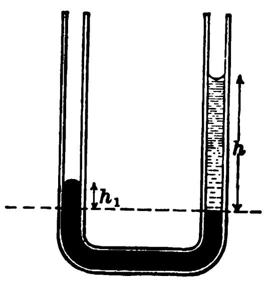


FIG. 82.

the two arms so as to coincide with the plane where the two liquids meet. All points of the liquid in both arms, which are in this plane, must have the same pressure, because they are all points in the same liquid, viz. the denser one. But if the vertical height of the free surface of one

liquid above this plane is h_1, the pressure in this plane due to it is $\rho' g h_1$ if ρ' is the density of that liquid; and, if the vertical height of the free surface of the other liquid above this same plane is h, the pressure due to it is $\rho g h$.

Hence $$\rho g h = \rho_1 g h_1,$$

or $$\rho : \rho_1 = h_1 : h \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Consequently, if the density of one liquid is known, that of the other may be determined by measuring h and h_1. The density of water at 4 ° C is 1 (see Art. 8); so that if water is used as one of the liquids, the density of any liquid which does not mix or react with it may be determined by this method.

If the two liquids do mix or react, a simple modification of the experiment will obviate the difficulty. Invert the U-tube and place the open ends in two vessels which contain the two liquids of density ρ and ρ_1. By an opening in the connecting tube at the top, exhaust some of the air in the tubes until its pressure is p. The liquids will rise in the tubes to vertical heights h and h_1 above their free surfaces. At the free surface of each liquid, in the open vessels, and therefore at these levels inside the tubes, the pressure is that of the atmosphere. But in the tube containing the liquid of density ρ, this pressure is due to a pressure, p, on the top of the column of liquid plus the pressure $\rho g h$ due to the liquid itself. Hence

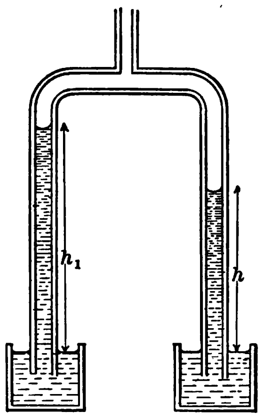


FIG. 83.

$$\text{Pressure of atmosphere} = p + \rho g h.$$

Similarly in the other tube

$$\text{Pressure of atmosphere} = p + \rho_1 g h_1.$$

Hence $$\rho g h = \rho_1 g h_1,$$

or $$\rho : \rho_1 = h_1 : h.$$

So that, if water at 4° C. is used as one of the liquids, or if any liquid having a known density is used, the density of any other liquid may be at once determined.

In both of these experiments the tubes which contain the liquids must be so large that there is no capillary effect (see Art. 96); but there is no other restriction as to their size or shape.

92. Archimedes' Principle. If a liquid is not flowing, any sensible portion of it is at rest, and may be considered as enclosed in a thin film so as to be separated from the rest of the liquid. The weight of this portion tends to make it move downward, but, since it does not move, there must be an equal and opposite force acting upward. This upward force is the resultant of all the pressures acting at each point of the film, perpendicular to the surface of the film. Consequently, the resultant of all these surface pressures on the film is a force vertically upward, equal to the weight of the liquid inside and passing through its centre of gravity. It is supposed, of course, that the bottom of this film is not on the bottom of the vessel; for, in that case, there would be no pressure up due to the liquid.

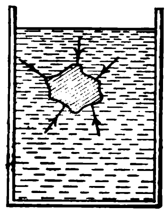


Fig. 84.

Let this portion of the liquid be so chosen that it has the same shape and size as some solid which is denser than the liquid, and which is not acted upon by it. Now, if the liquid inside the film be removed, and its place taken by the solid, the surface pressures over the film have not been changed, because the film has not changed in the least. So, while the force down on this solid is

its weight, there is also a force up equal to the weight of the liquid which has been displaced, if the bottom of the solid does not rest on the bottom of the vessel. This fact is known as "Archimedes' Principle." So, if the solid is hung from one pan of a chemical balance, it will be noticed that, when it is immersed in the liquid, there is less force pulling it down. This "loss in weight" is due to the force upward which equals the weight of the liquid displaced.

93. Measurement of Density of a Solid. Let the volume of the solid be v, its density ρ; then its mass is ρv, and its weight $\rho v g$. When it is totally immersed in the liquid, the volume of the liquid which is displaced is also v; and, if its density is ρ', its weight is $\rho' v g$. This is the force upward, if the solid does not rest on the bottom of the vessel; and it may be measured by weighing the solid first in air (strictly, in a vacuum), then when immersed in the liquid. The weight in air (strictly, in a vacuum) is $\rho v g$; the difference in the two weighings, or the "loss in weight," is $\rho' v g$. The ratio of these two quantities is ρ / ρ'; and so, if the density of the liquid is known, that of any solid denser than it, and which is not acted upon by it, may be determined. Water at 4° C. is generally the liquid used, because its density is 1. If a solid has a density less than that of the liquid, it may be immersed in the liquid by hanging some very heavy solid below it, and then keeping this last solid immersed in the liquid during both weighings.

This principle also gives a method by which the densities of two liquids may be compared. If the solid is immersed in turn in the two, its loss in weight in them will be $\rho_1 g v$ and $\rho_2 g v$ respectively; which gives at once the ratio of the two densities.

Since, then, a liquid produces an upward force upon any immersed solid, which does not rest on the bottom, there must be an equal reaction of the solid upon the liquid.

So that, if the vessel containing the liquid stands on a platform-balance, there will be an additional force down, which will be registered on the balance, when a solid is held immersed in it, but not touching the bottom. This additional force equals the weight of the liquid displaced. So a platform-balance may be used, as well as a chemical balance, for the measurement of densities.

Other instruments, such as Hydrometers, Jolly's Spring Balance, etc., have been devised and are in daily use for these same measurements.

94. Floating Bodies. If a body floats on the surface of a liquid, it is in equilibrium; and so the forces in any direction must be balanced. The force down is its weight, and it acts in a line passing through the centre of gravity of the body. The force up equals the weight of the liquid displaced, and passes through the centre of gravity of this liquid.

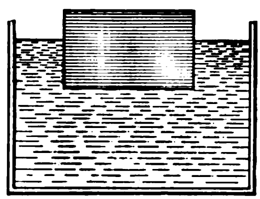


FIG. 85.

So, when a body is floating, it displaces its own weight of the liquid; and the centre of gravity of the body must lie in the same vertical line as that of the liquid displaced.

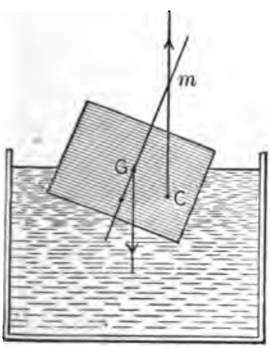


FIG. 86.

This equilibrium may be stable, unstable, or neutral. The question must be investigated by considering what happens when the body is slightly tipped one side. When the body is in equilibrium, the two centres of gravity, that of the body and that of the displaced liquid, must lie in the same vertical line, which is called the "axis." Draw this line in the body, and consider it fixed there. Now, if the body is

tipped, the axis will be inclined to the vertical; it will still pass through *G*, the centre of gravity of the body, because both it and the axis are fixed in the body; but the centre of gravity of the displaced liquid, *C*, will now lie to one side. So there is a couple acting on the body, made up of the two equal forces, one down through *G*, the other up through *C*. Let a line drawn vertically through *C* meet the axis in a point *m*, called the "metacentre." If *m* lies above *G* on the axis, it is evident that the couple acting on the body tends to restore the axis to its vertical position; and so the equilibrium was stable. While, if *m* is below *G* on the axis, the couple will tend to make the body tip farther; and the equilibrium was unstable. The position of the metacentre depends upon the shape and construction of the body and upon the way it is displaced; and, in making calculations to determine it, the tipping is assumed to be very slight. Any ordinary boat has two principal metacentres, — one for a rolling motion, the other for a pitching one.

95. Work Required to Change the Volume of a Liquid. As a consequence of the equality of pressure throughout a liquid, it is not difficult to calculate how much work is required to increase the volume of a liquid. (This increase may be produced by forcing in more liquid or by raising the temperature.) Imagine the liquid enclosed in a cylinder, into one end of which is fitted a plane movable piston. Let the area of this piston be A, and the pressure of the liquid on it p. Then the force is $p A$; and, if the piston is driven out a distance l owing to the increase in volume, the work done is $p A l$, if the pressure remains constant. But $A l$ is the increase in volume; and so the work done at constant pressure is the product of the pressure and the change in volume. (If during the change in volume the pressure

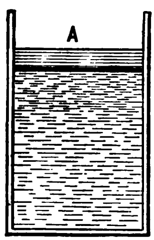


FIG. 87.

changes uniformly, the average pressure is the mean of the initial and final values; and the work done is the product of the average pressure and the change in volume.)

$$\text{work} = p\,(v_2 - v_1) \quad . \quad . \quad . \quad . \quad . \quad (11)$$

96. Liquids in Motion. If a liquid which had no viscosity was allowed to escape through an opening in a thin wall at the depth h below its free surface, its speed of efflux would be

$$s = \sqrt{2\,g\,h} \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

Because, owing to the fact that there is in this case no resistance to the motion of one portion of the liquid over another, a particle of the liquid as it escapes will have the same speed as if it had fallen freely from the free surface through the height h. (See Art. 18.) The pressure at a point just outside the opening is that of the atmosphere, while that just inside is greater than this by an amount $\rho\,g\,h$. So if p is the difference of pressure outside and in at the opening,

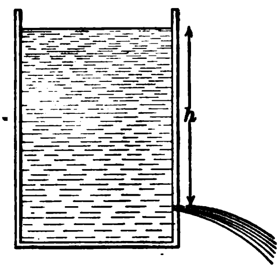


FIG. 88.

$$p = \rho\,g\,h;\ \text{and so}\ s = \sqrt{\frac{2\,p}{\rho}} \quad . \quad . \quad . \quad (13)$$

All actual liquids are viscous; and so the observed speeds of efflux are always less than this. If the liquid escapes through a long tube, it is found by experiment that the material of the tube has no effect on the rate of efflux. This proves that there is a layer of the liquid formed over the inner wall of the tube, and that this does not move, the friction being then between the layers of the liquid, and depending only upon the viscosity of the liquid.

If two liquids are placed directly in contact, there is a gradual diffusion of one into the other, which goes on entirely independently of their relative densities. In time the mixture would be uniform.

Certain membranes and porous materials such as parchment, unglazed earthenware, etc. have the property of being permeable for some substances and not for others. This general phenomenon of the passage of substances though semi-permeable membranes has received the name "osmosis." If a miscellaneous mixture of a great many substances is placed in a parchment dish and floated on water, it is noticed that after a certain time some of the substances have passed through into the water, while others never will.

If a solution of cane-sugar in water is enclosed in a tube over one end of which is a parchment cap, and if this end is dipped into pure water, it is observed that the water passes through the parchment into the tube. So that the level of the liquid inside the tube becomes higher than it is outside by an amount h, when the flow stops; that is, it takes a pressure $\rho g h$ to overcome the tendency of the water to come in. This tendency is caused by the sugar in solution; and the pressure $\rho g h$ is called the "osmotic" pressure of the solution.

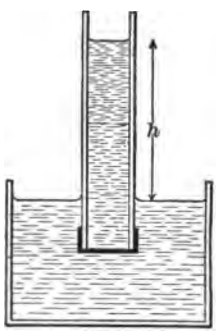


FIG. 89.

In the flow of a liquid through a tube of any shape, if the speed is greater at one point than at another, the pressure at the first point must be less than at the other; because to produce an *increase* of speed a force is necessary, that is, a *fall* in pressure. This is the explanation of a ball in a fountain remaining poised, of the so-called "ball-nozzle," and several other phenomena.

97. Capillarity. The property of a liquid which distinguishes it from a gas is that it has a definite size, and so is separated from other liquids or gases by a definite surface unless there is diffusion. As a consequence of a liquid having this surface, it has certain characteristics, and obeys certain laws.

It is at once seen that a particle of a liquid in its free surface is in a different condition from a particle in the interior; for the latter is influenced by particles of the liquid on all sides of it, while the one in the surface is influenced by only half as many. The fundamental property of a liquid surface is its tendency to contract and so become as small as possible. This is proved by a great many experiments. A mixture of water and alcohol can be made which has the same density as olive oil. So, if a drop of oil is put in a vessel of this mixture, it will stand anywhere in the water entirely uninfluenced by the force due to the earth. Now, it will be found that this drop will have a spherical form; and a sphere has the least surface for a given volume. Ordinary drops of water and other liquids are not spherical, because their shape is changed owing to the earth's force of gravitation.

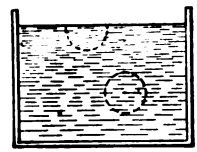


FIG. 90.

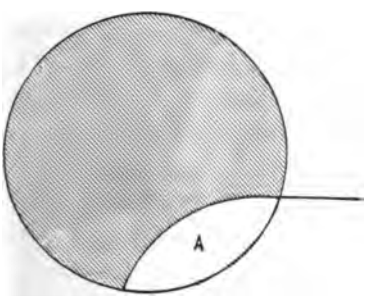


FIG. 91.

Again, if a film of a soap-mixture stretched across a circular wire frame is broken at any point, it will instantly contract. The following experiment illustrates the same fact. Fasten a string to one point of the wire frame, and, holding the other end of the string, dip the frame in a soap-mixture so as to make over it a film,

in which the string lies. Then break the film on one side of the string, and the film will contract until the stretched string forms the arc of a circle; for then the film is as small as it can be under the conditions. If the string is loosened slightly, the film contracts still more; and, if it is desired to stretch the film, force must be exerted on the string.

So it is proved that there are forces acting in any liquid film, tending to make it contract. If a line is imagined drawn on the surface of a liquid, there is a certain force acting across it holding together the portions of the surface on its two sides. The force acting across one centimetre of any line in a surface is called the "surface-tension," T.

Thus, consider any line 1 cm. long in the film, as shown; there are forces, T, in both directions perpendicular to the line. Since they are equal, the line does not move; but, if the surface-tension on one side is weakened in any way, then the line will move towards the other side. This may be shown by heating the film on one side or by putting some impurity in it; for the surface-tension is diminished by both these causes. This explains the rapid motion given to a bit of camphor or sodium when tossed upon a surface of water. The solid dissolves unequally at various points; so the surface is rendered impure at some points faster than at others; and the surface-tension being thus weakened more at these points, there is motion in the opposite direction.

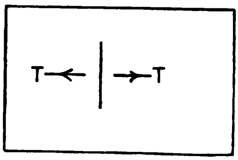


FIG. 92.

The surface-tension of a pure liquid in contact with any definite substance is constant for a given temperature; for, when the surface is increased, there is not a *stretching* of the surface, but the formation of more surface of the same liquid. For this reason a vertical film of a pure liquid

cannot be made here on the surface of the earth; for, owing to the weight of the liquid in a vertical film, the forces must be greater at the top than at the bottom.

Owing to this tendency of a surface to contract, a curved surface tends always to have a smaller radius; and so there is a pressure produced toward the centre of curvature. Thus, a bubble will contract until stopped by the pressure of the gas inside. A drop contracts until stopped by the pressure of the liquid. It is not difficult to find a relation between this pressure of the liquid in a drop and the surface-tension. Imagine the drop cut into two hemispheres, and then replaced. They will be held together by the forces acting across the equator where the cut is made in the surface. If r is the radius of the sphere, the length of the equator is $2\pi r$; and, as T is the force across one centimetre, the entire force holding the two hemispheres together is $2\pi r T$. As a consequence of this force, the hemispheres are pressed together until a pressure p is reached which counterbalances the force. The area of the equatorial section is πr^2; and, as p is the force per unit area, the entire reaction of the two hemispheres against each other is $\pi r^2 p$. Hence, since action and reaction are equal but opposite,

$$2\pi r T = \pi r^2 p;$$

and so

$$p = 2T/r \quad . \quad . \quad . \quad . \quad . \quad (14)$$

Therefore a surface whose tension is T and whose radius of curvature is r produces a pressure towards the centre of curvature equal to $2T/r$.

If the surface, instead of being a sphere, had been a cylinder of radius, r, the pressure would have been

$$p = T/r \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

If certain solids are immersed in certain liquids and then taken out, they are found to be covered with a film

of the liquid, and are said to be "wet" by it. Thus, glass is wet by water. Other substances are not wet; e. g. glass is not wet by mercury under ordinary conditions.

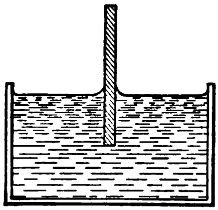


FIG. 93.

If a plate of glass, *previously moistened*, is dipped in water, the film of the water is at first rectangular, being part on the glass and part on the surface of the water. But this film tends to contract; and so the corner is rounded off, and the water is said to rise against the glass. The contraction of the film will continue until stopped by the weight of the liquid raised. Let a glass tube of small bore, *previously moistened inside*, be dipped into a vessel of water. The water will rise against the glass, and in the interior the water will rise to a considerable height if the bore is small enough. For the liquid film inside the tube is over the walls of the tube and over the water where the tube enters the surface; that is, it is like the finger of a glove. This surface tends to contract; and it can do so if the liquid surface inside the tube rises; so it will continue to rise until stopped by the weight of the liquid raised. Since the water wets the glass, the surface of the liquid inside the tube is spherical, if the tube is circular, and the radius of the sphere equals the radius of the tube. The pressure at a point of the level surface of the water outside the tube is that due to the atmosphere; so it must be this also at the same level inside the tube. But the pressure there is due to three causes, — the pressure of the atmosphere on top of

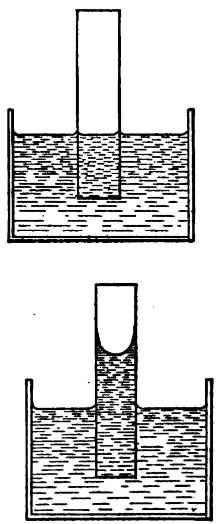


FIG. 94.

the liquid column, the pressure upward due to the curved surface, the pressure due to the height of the column. Call this vertical height h, and the radius of the tube at the point where the curved surface is, r. Then,

$$\text{Atmospheric pressure} = \text{atmospheric pressure} - 2\,T/r + \rho g h.$$

Hence $$\rho g h = 2\,T/r;$$

and so $$h = \frac{2\,T}{\rho g r} \quad \ldots\ldots \quad (16)$$

Consequently the vertical height of the column varies inversely as the radius of the tube at the point where the curved surface is, but is independent of the shape or size of the tube elsewhere, because the pressure in a liquid depends only on the depth of the liquid, not on the shape of the containing vessel. So, if h, ρ, g, r are measured, T may be calculated. T will vary with the temperature, always decreasing as the temperature rises; it will also vary with the substance with which the liquid is in contact. These phenomena due to the surface-tension were first observed in the rise of liquids in small tubes, and so are often called "capillary" phenomena, because the bore of the tubes is comparable to the size of a hair, the Latin word for hair being "capillus."

In identically the same manner as this, it may be explained why, if a glass tube is dipped into a vessel of mercury, the mercury inside the tube sinks.

Since in any curved surface $p = 2\,T/r$, it would require an infinite pressure to make a surface of radius O; and the smaller the radius, the greater the pressure. This is why the formation of drops or bubbles is greatly facilitated by the presence of points or nuclei; for, as they begin to form, the curvature of the surface is that of the point or nucleus, and so is of finite dimensions. This has a most important bearing on the condensation of any vapor, and also on the boiling of liquids.

A soap-bubble has two films of nearly the same radius: so the pressure of the gas inside must be $4\ T/r$.

When two solids which are wet by a liquid are dipping in the liquid and are brought so near that the liquid rises between them, there will be a motion of the two solids towards each other, because in between the two plates at points above the level of the liquid outside the pressure is less than the atmospheric pressure; and so the solids are pressed together by the air outside. Similarly, it may be shown that, if the two solids are not wet by the liquid, they will also be pressed together when they are dipping in the liquid. But two solids, one of which is wet by the liquid, the other not, are pushed apart, if they are dipping in the liquid near each other.

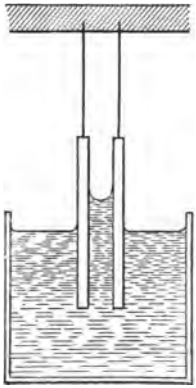


FIG. 95.

If a drop of a liquid is placed in a conical tube, it may move in one direction or the other, depending upon the different curvatures of the surfaces at its two ends.

SPECIAL PROPERTIES OF GASES

98. Introduction. The distinguishing properties of a gas are: 1. Any force, no matter how small, can cause one portion to move over another; 2. It has no definite size of its own, but fills the space enclosing it. As a consequence of the first property, all the laws of pressure which were derived for a liquid hold equally well for a gas; but, as the gas has no free surface, and can have its volume changed most easily, it has certain new properties in connection with its volume; 3. The molecules of a gas are in rapid motion in all directions, on the kinetic theory.

99. Properties of a Gas at Rest. Without further proof, then, the following statements are true:—

1. The pressure due to a gas is always perpendicular to any surface in contact with it.

2. The pressure at any point in a gas is the same in all directions.

3. The pressure at any point in a gas is due to two causes: (*a*) the pressure of the enclosing vessel; (*b*) the pressure due to the weight of the gas above the point.

4. Archimedes' Principle applies perfectly to gases. That is, a solid immersed in a gas is buoyed up by a force which equals the weight of the gas displaced. So, in weighing solids in the air, a correction must be made for the buoyancy of the atmosphere.

100. Pressure due to Weight of Atmosphere. Unless a column of gas is very high, it produces a comparatively small pressure; for, if the density of the gas is ρ, the difference in pressure between two points at a vertical height h apart is $\rho g h$. ρ is very small for gases; so h must be large, if this pressure is to be sensible. An illustration of a high column of a gas is afforded by the earth's atmosphere, which does produce a large pressure. It may be measured by balancing it against a column of a liquid. Completely fill with mercury a glass tube, which is closed at one end, of rather large cross-section, and about 80 cm. long; invert it, allowing no air to enter; and place the open end in a basin of mercury. It will be observed that the mercury will continue to stand in the tube up to a certain height above the free surface in the basin. The pressure must be the same at the free surface of the mercury as at the same level in the tube. The pressure outside is the atmospheric pressure due to the air; that inside is $\rho g h$, where ρ is the density of the mercury and h is the height of the column. This

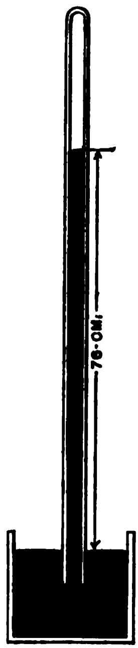


FIG. 96.

is the entire pressure, because there is no gas pressing down on the top of the column, there being a vacuum above the column, except for the slight amount of mercury vapor which may arise from the liquid mercury. Hence

$$\text{atmospheric pressure} = \rho g h \quad . \quad . \quad . \quad (17)$$

h varies from hour to hour, but is always in the neighborhood of 76 cm. for mercury. Such an instrument as this, which measures the pressure of the atmosphere, is called a barometer. The height to which the liquid stands does not depend on the shape or size of the tube; for it is the *pressure* which determines the height. At different temperatures, however, the density of the liquid is different; and so, even though the pressure is unchanged, the height h may vary. Consequently, to compare pressures by the heights of the barometric column, these must be given in terms of some standard liquid which does not change. This is done by "correcting" the actual reading on the barometer in a manner explained in Heat (Art. 175), by calculating, from the observed temperature and height, that height to which the column would have risen if the density of the liquid were that which it has at 0° C.

There are various forms of barometers, and various liquids are used; but in each the pressure is given by $\rho g h$, where ρ is the density of the liquid, and h is the height from the free surface to the top of the column. Thus, if the liquid is mercury, whose density is 13.6, and if the height is 76 cm., the atmospheric pressure is $13.6 \times 980 \times 76 = 1013000$ dynes per sq. cm.

There are innumerable illustrations of this pressure of the air. If two metal hemispheres are carefully fitted together, and the air is pumped out from the interior, a great force is necessary to pull them apart, because the pressure of the air holds them together. If a narrow goblet is filled with water, then covered with a card and carefully inverted, the water will not escape.

101. Pumps. The ordinary lifting-pump is also an illustration of the pressure of the air. It consists of a cylinder in which moves a close-fitting piston, and which is connected at its lower end by a pipe with the well or supply of water; in the piston at B and at the top of the pipe at A there are valves opening upward. When the piston is being raised, its valve, B, is closed, and the pressure of the air below the piston is so diminished that the atmospheric pressure on the free surface in the well forces water up through the pipe and the valve A into the cylinder. Now, if the piston is forced down, the valve A closes and that at B opens, allowing the water to pass through above the piston. So, when the piston is again raised, this water is lifted and runs out at the top of the pump, while the cylinder below the piston becomes again full of water. And so the process continues. Of course the pump cannot work if the pipe from the well is so long that the atmospheric pressure cannot lift the water that high. This height is given by the condition that $\rho g h$ for water equals the atmospheric pressure. If this is 76 cm. of mercury,

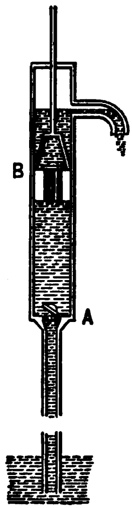


FIG. 97.

$$1 \times g \times h = 13.6 \times g \times 76.$$

Hence the limiting height h is

$$h = 13.6 \times 76 = 1033.6 \text{ cm.}$$

In practice this height can never be reached in a pump, owing to leaks around the piston, which allow some air to remain below it.

102. Siphon. Another illustration of atmospheric pressure is given in the "siphon," which consists of a tube bent into a U with one arm longer than the other, and which is used to make a liquid flow out of a basin over its

edge. The tube is filled with the liquid, and is then placed with its shorter arm dipping in a basin of the liquid, and its longer arm projecting down outside the basin. To do this, the lower end, B, of the longer arm must be kept closed; and the pressure at a point, A, in this arm at the level of the free surface in the basin will be at the atmospheric pressure, because of the equality of pressure at the same level in connecting tubes of the same liquid. So, if now the end B is opened, the pressure of the liquid between B and A will be in excess of the atmospheric pressure outside, and the liquid will start to run out; but, as it does so, the pressure in the long arm is diminished, and the atmospheric pressure on the surface of the liquid in the basin forces over some liquid, and so the stream is kept up. The vertical height of the top of the siphon above the free surface in the basin must not be greater than corresponds to the atmospheric pressure, or the liquid will not flow out. Hence there would be no flow at all (at least from this cause) if the basin and siphon were placed in a vacuum.

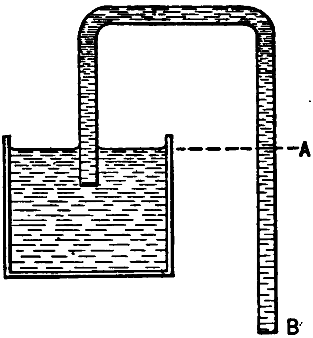


FIG. 98.

103. Measurement of Pressure of a Gas. If the pressure of the atmosphere is known, the pressure of any gas enclosed in a vessel may be measured. Insert in the side of the vessel a bent tube open at both ends, which contains some liquid, e. g. mercury. The liquid will in general stand at different heights in the two tubes; and, if the difference in the two levels is h, the pressure of the gas inside is different from that of the atmosphere by an amount $\rho g h$, where ρ is the density of the liquid. If the liquid is forced up in the open tube so as to stand higher

there than in the arm connected with the gas, the pressure of the gas is greater than that of the atmosphere by the amount $\rho g h$; whereas, if the contrary is true, the pressure of the gas is less than that of the atmosphere by $\rho g h$. Such an apparatus for measuring pressures is called an open manometer.

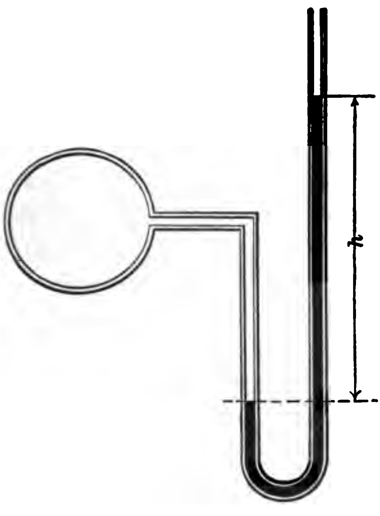


FIG. 99.

FIG. 100.

If a gas is enclosed in the space above a column of mercury in a barometer tube, its pressure will force the mercury down a certain distance; and, if the pressure of the atmosphere is known, that of the gas can be measured at once. Let the column of mercury stand at the height h. Then, equating the pressures outside and inside the tube at the level of the free surface, atmospheric pressure $= \rho g h +$ pressure of gas above column, and so pressure of gas $=$ atmospheric pressure $- \rho g h$.

104. Work Necessary to Change the Volume of a Gas. As a further consequence of the equality of pressure throughout a gas, it is at once evident that, as in the case of a liquid, the work done in changing the volume while the pressure is constant is the product of the pressure and the amount of the change in volume (see Art. 95).

105. Gases in Motion. Just as a liquid, escaping through a thin opening under a pressure p, has a speed $s = \sqrt{\frac{2p}{\rho}}$, so a gas under the same conditions obeys the same law. The pressure, p, which produces the flow of the gas is the difference of pressure of that particular gas on the two sides of the opening; and the density, ρ, is that of the gas. By measuring, then, the quantities of different gases which escape in a given time under the same conditions, it is possible to determine the ratio of the densities of any two gases. (Another method of determining densities would be to weigh a large bulb, of known volume v, first empty and then filled with the gas. If the difference in weight is m grams, the density of the gas under the existing conditions of temperature and pressure is $\frac{m}{v}$.) Gases, like liquids, diffuse into each other; so that in time the mixture becomes uniform. Gases also pass through many solids, and are often absorbed by solids and liquids. Since the rate of escape through small openings varies inversely as the square root of the density, dense gases pass through porous bodies much more slowly than do rare ones.

If a gas is flowing through any irregular space, the pressure is least where the velocity is greatest, just as is the case of a liquid. This explains why, if a jet of air is blown over the end of a tube the other end of which dips into a liquid, the liquid rises in the tube and may reach the top and be blown away, as in an "atomizer."

106. Properties of Gases as Distinct from Liquids. The particular property of a gas which distinguishes it from a

liquid is the fact that it takes the size of the containing vessel. Consequently a gas is very compressible, because the volume of the containing vessel may be changed as one wishes. The particles of a gas, i. e. its smallest portions, are also moving about with a great freedom of motion, even more than is the case with a liquid. As a consequence of these facts, there are certain so-called laws of gases, which have been found to be true by careful experiments.

107. Dalton's Law. This states that, if several gases which do not react on each other are put inside a certain vessel, the mixture becomes uniform throughout, and the pressure at any point is the sum of the pressures which each gas would produce if it occupied the vessel by itself. This is equivalent to saying that the particles of each gas are moving rapidly in all directions, and that their actual size is so small compared with the volume of the vessel that they do not in any way influence each other.

108. Boyle's Law. This states that, if the temperature is kept unchanged, but the density of the gas altered, the pressure is always proportional to the density. If p is pressure, ρ density, v volume, m mass, Boyle's law may be written: at constant temperature,

$$p = k\rho, \quad \text{or} \quad pv = km \quad . \quad . \quad . \quad . \quad (18)$$

k is a constant for a given gas and definite temperature.

By most elaborate experiments at high pressures it is found that Boyle's law is not rigidly true, but that the ratio p/ρ slightly increases for enormous pressures.

If the volume of a given mass of gas is known at any pressure, the pressure which corresponds to any other volume may be calculated if the temperature has not changed. For $pv = p_1 v_1$. An apparatus called a "closed" manometer has been devised on this principle to measure pressures. A strong glass tube containing air is fastened into an iron box full of mercury. This box is connected

with the space where the pressure is to be determined. By a preliminary experiment the volume, v, of the air in the tube is measured when the pressure on it is that due to the atmosphere; and when the mercury is subjected to the unknown pressure, p_1, let the air have its volume changed to v_1. Hence, if the temperature has not changed,

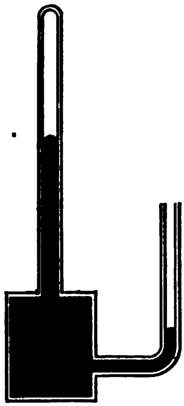


FIG. 101.

$$p_1 v_1 = v \times \text{atmospheric pressure},$$

and so p_1 may be calculated.

109. Elasticity of a Gas. The only coefficient of elasticity for a gas is the one corresponding to a change in volume; and it may in certain cases be easily calculated. For, the stress is the increase in pressure; and the strain is the ratio of the decrease in volume to the original volume, the changes being very small. That is, if p_1, v_1, are the original pressure and volume, and p_2, v_2, the resulting ones due to the change,

$$\text{the coefficient of elasticity} = \frac{p_2 - p_1}{\frac{v_1 - v_2}{v_1}} = \frac{v_1 (p_2 - p_1)}{v_1 - v_2} \quad (19)$$

If the compression takes place so slowly that there is no change in temperature, Boyle's law may be applied: $p_1, v_1 = p_2, v_2$. Hence

$$\frac{v_1 (p_2 - p_1)}{v_1 - v_2} = \frac{p_2 (v_1 - v_2)}{v_1 - v_2} = p_2 \quad . \quad . \quad . \quad (20)$$

But, since the change in the pressure has been presupposed very small, it may be stated that the coefficient of elasticity of a gas at constant temperature numerically equals the pressure of the gas at that instant.

If the compression takes place rapidly, there is always an increase in temperature; and so Boyle's law cannot

be applied. However, if the gas is compressed extremely rapidly, so that there is no time for it to give out heat to the surrounding bodies, it may be proved that the coefficient of elasticity of the gas is a certain constant times the pressure, γp, where γ has the value for air, hydrogen, oxygen, and many other gases of about 1.40. The nature of γ will be explained later, under the subject of "Specific Heat" (Art. 182).

110. Air-pumps. Just as water-pumps are used to draw water out of one vessel and pour it into another, so pumps may be constructed to exhaust a gas from a vessel which contains it. Such pumps are called air-pumps, and are of three general types.

1. Mechanical pump. This is in principle identical with the ordinary liquid lifting-pump, with three valves, A at the bottom of the cylinder, C at the top, and B in the piston. The vessel from which the gas is to be exhausted is connected in some way with the pipe leading out of the bottom of the cylinder. As the piston is raised, the valve B in it is closed; the gas above it is forced out into the air through the valve C; and, the valve A being opened, the gas from the vessel expands into the cylinder below the piston. If the piston is now forced down, the valves A and C are closed; and, the valve B in the piston being opened, the gas below the piston passes through into the space above. So, as the process continues, more and more of the gas is exhausted. The three valves will not open and close of themselves, as they would in a water-pump, but are made to do so by automatic mechanisms connected with the piston.

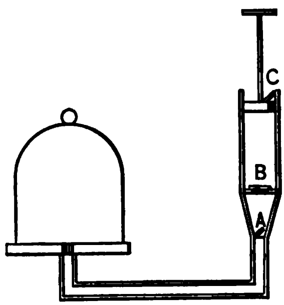


FIG. 102.

2. Sprengel pump. The action of this pump consists in having drops of mercury (or some other liquid) so fall as to trap the gas between them, and thus carry it away. There is an elongated glass bulb, into the upper end of which is joined a tube provided with a stop-cock and funnel so that the mercury may be thus poured in; into the lower end of the bulb is joined a glass tube of narrow bore and at least 80 cm. long; into the side of the bulb is joined a connection with the space to be exhausted. The tubes at the top and bottom of the bulb are so arranged that, as the mercury-drops break off and fall, they hit the opening of the lower tube and pass down it in the form of short cylinders. The space between these cylinders thus formed is occupied by small amounts of the gas, drawn in from the connected vessel; and so these drops act like a succession of small pistons forcing out the gas. The lower end of the long tube may dip into a basin of mercury; and the gas will bubble out at the surface. As the exhaustion continues, the mercury will rise in the long tube, and will finally stand at the barometric height when the vacuum is as complete as it can be made.

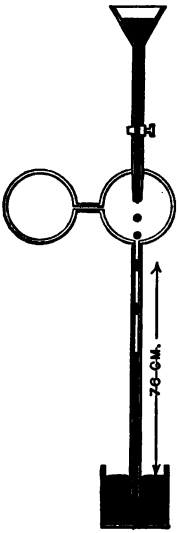


FIG. 103.

3. Geissler-Toepler pump. In this pump there is a large bulb to which are joined two tubes, — one at the top, the other at the bottom. The lower one is at least 80 cm. long, and is connected at its lower end to a large vessel of mercury by means of a long rubber tube. The upper tube is bent over into a vertical direction downward, and dips into a basin of mercury. Around the large bulb there is a glass-tube branch connecting the upper and lower tubes just as they leave the bulb ; and into this branch is joined

a long vertical tube leading to the vessel which is to be exhausted.

If the large vessel of mercury is now raised, as it can be owing to the flexible rubber tubing, the mercury will rise into the bulb and the connecting tubes, shutting off connection with the vessel to be exhausted, and will drive

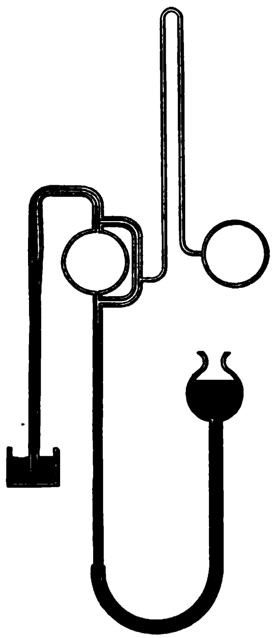


Fig. 104.

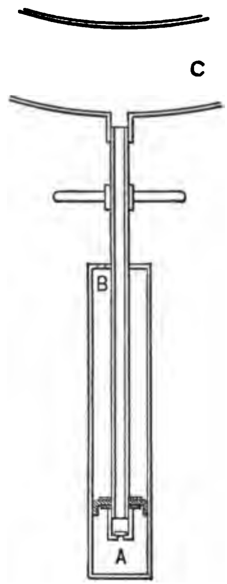


Fig. 105.

out all the gas in the bulb through the tube in the top, so that it will bubble out through the mercury in the basin at its end. If, now, the movable vessel of mercury is lowered, no air can enter through the tube at the top of the bulb, because it is "sealed" by the mercury in the basin, which will rise in the tube; but, as soon as the mercury falls below the opening to the long vertical tube, the gas in the vessel to be exhausted will expand and fill the bulb

and the connecting tubes. When the movable vessel of mercury is again raised, it drives out the gas in the bulb; and, as the process continues, the exhaustion of the vessel rapidly proceeds. The tube leading from the top of the bulb around to the basin of mercury must be at least 80 cm. high; and the long vertical tube leading to the vessel to be exhausted must be still longer.

111. Compression Pumps. One often desires to compress a gas in a tube or cylinder, and for this purpose so-called compression pumps are used. They are, in principle, nothing but mechanical air-pumps with their valves arranged so as to open in the opposite directions. A simple form of compression-pump is the ordinary bicycle pump, a section of which is shown. (See Fig. 105.) Around the inner tube, *A*, is a leather "washer," which allows air to pass by it when the outer tube is drawn away from the tire. The outer tube thus becomes filled with air; and, if it is now pushed in, the washer does not allow the air to escape around it, and so the air is driven into the tire through the tube *A*. There is, of course, an opening, *B*, in the outer tube, joining it to the open air so as to allow the air to enter at each stroke of the pump.

TABLE I

DENSITIES

SOLIDS

Aluminum . . .	2.58	Iron	7.86
Brass . (about)	8.4	Lead	11.3
Copper	8.92	Platinum . . .	21.50
Diamond . . .	3.52	Silver	10.53
Glass, common .	2.6	Tin	7.29
" heavy flint .	3.7	Zinc	7.15
Ice at 0° C. . .	0.9167		

Liquids

Alcohol at 20° C. .	0.789	Mercury	13.55
Ethyl ether at 0° C.	0.735	Water at 4° C. . .	1.000

Water at other temperatures, see below.

Gases at 0° C. and 76 cm. of Mercury

Air, dry . . .	0.001293	"Helium" . .	0.00021
"Argon" . .	0.00170	Hydrogen . .	0.0000895
Carbon dioxide	0.001965	Nitrogen . .	0.001254
Chlorine . . .	0.00317	Oxygen . . .	0.001429

Water at different Temperatures

0° C. . . .	0.999878	16° C. . .	0.999004
1° . . .	0.999933	17° . .	0.998839
2° . . .	0.999972	18° . .	0.998663
3° . . .	0.999993	19° . .	0.998475
4° . . .	1.000000	20° . .	0.998272
5° . . .	0.999992	21° . .	0.998065
6° . . .	0.999969	22° . .	0.997849
7° . . .	0.999933	23° . .	0.997623
8° . . .	0.999882	24° . .	0.997386
9° . . .	0.999819	25° . .	0.997140
10° . . .	0.999739	26° . .	0.99686
11° . . .	0.999650	27° . .	0.99659
12° . . .	0.999544	28° . .	0.99632
13° . . .	0.999430	29° . .	0.99600
14° . . .	0.999297	30° . .	0.99577
15° . . .	0.999154	31° . .	0.99547

BOOK II

SOUND

BOOK II

SOUND

INTRODUCTION

To any one with the sense of hearing, the name "sound" conveys a definite idea of a particular sensation. In Physics the theory of sound is the study of the nature of this sensation and of the exact conditions under which it is produced.

112. It is a matter of every-day experience that sounds are always caused by vibrating bodies; e. g., a piano-string, the air in an organ-pipe, an explosion or sudden blow. It is also probably well known to every one that, if the vibrating, i. e. "sounding," body is far away, some time elapses between the vibration and the perception of the sensation. Thus it is noticed that there is an interval of time between the flash of a pistol and the sound of its report; another illustration is the interval between the flash of lightning and the consequent thunder. It may also be shown that, if there is no material medium between the sounding body and the ear, no sound is heard. For, if a bell is rung inside a vacuum, the ear hears no sound. This proves that the presence of matter is essential for the propagation of the disturbance that produces sound.

Further, it is important to emphasize the fact that sound is a sensation perceived in the ear. The sound is not emitted by the vibrating body, but is produced by it.

CHAPTER I

VIBRATIONS

113. Detection of Vibrations. That sounding bodies of all kinds are in vibration admits of immediate proof. Generally, the sense of touch is sufficient to convince one of the fact; for the vibration may be felt. This is true in the case of solid bodies, such as a stretched string, a tuning-fork, or a bell. When columns of air in organ-pipes are causing sounds, the air may be easily proved to be vibrating by placing at different points in the pipe membranes carrying some light powder. The powder will in certain places be violently shaken. Similar means may be used to detect the vibration in the case of sounding columns of a liquid.

Nature of the Vibrations. The exact nature of the vibrations may also be studied by suitable methods. One of the

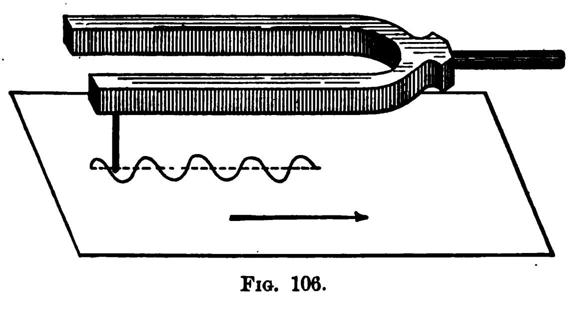


Fig. 106.

simplest is to place a stiff bristle upon the body, and then while it is vibrating to draw under it a piece of smoked glass, so that the bristle traces on the glass a path which corresponds to the vibration of the body and the motion

of the glass. Such a method applied to a tuning-fork is shown in the figure.

Another method is to study the effect on the air in the neighborhood of the sounding body by placing near the body a stretched membrane to which is attached a bristle; and this can record the vibration of the air on a smoked plate.

(By having the glass plate attached to a second vibrating body, the combined motion of the two bodies may be registered and studied. This may also be done by suitable optical methods.)

In this way it is proved that there are two classes of sounding bodies. One class makes a series of periodic vibrations, the other does not. A periodic vibration is a motion such that at regular intervals of time the identical process is repeated. This interval is called the "period" of the vibration; and the number of complete vibrations in one second is called the "frequency." If T is the period, and n the frequency, it is obvious that $T = 1 / n$. Illustrations of this motion are the vibrations of tuning-forks, piano-strings, and all musical instruments.

Illustrations of non-periodic vibrations are the motions produced by a series of sudden blows, like a hammer on an anvil, a carriage rolling over a rough pavement, the tearing of a piece of paper, and all so-called noises. It may be shown that noises are due to a combination of a great number of periodic motions of nearly the same periods, but which last only an extremely short time.

The "amplitude" of any vibration is the extent of the motion between the extreme limits; and it is evident that the greater the amplitude is, so much greater is the energy of the vibration.

114. Periodic Vibrations. The simplest type of periodic vibration is harmonic motion (see Art. 21); but illustrations of this are by no means common among musical instruments. A tuning-fork or a stretched string may be

made to vibrate with harmonic motion; but it is not easy. In general, the motion is much more complicated. It may be proved, however, that any periodic motion, produced in any way, is equivalent in every respect to the combination of a number of harmonic motions of suitable periods and amplitudes. If the general periodic motion has the period T, the harmonic components must have the periods T, $T/2$, $T/3$, $T/4$, etc. The component with the period T is called the "fundamental" vibration; the others are sometimes called the "upper partial" vibrations. (This statement is known as Fourier's Theorem; and it admits of rigid mathematical demonstration.)

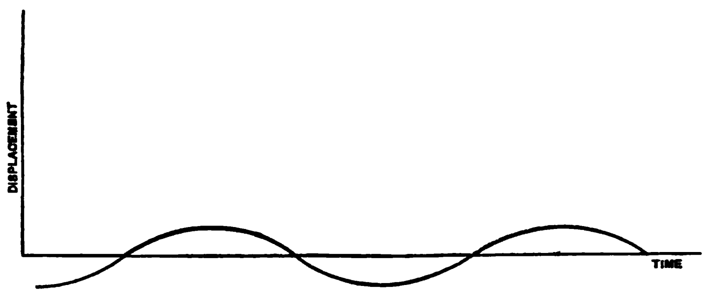


FIG. 107.

This question of the combination of harmonic vibrations so as to produce any periodic motion may be best illustrated graphically. Lay off along a horizontal line points corresponding to equal intervals of time; and at each point of this line erect a perpendicular line equal in length to the displacement at that instant of the vibrating body from its position of equilibrium. If the motion is harmonic, the curve which is the locus of the ends of these perpendicular lines will be as shown. (This curve is called a "sine-curve.") Any other motion will have its corresponding curve; and the combination of various vibrations may be shown by simply adding algebraically their curves. Thus, if there are two harmonic vibrations, whose periods are T and $T/2$

their curves are as shown; and their combination is given by simply adding the displacements which correspond to the same instants of time. The resulting curve will, of course, depend upon how much one vibration lags behind the other; that is, upon whether they begin together or at different instants. The curve shown corresponds to the case when the two vibrations begin at the same time; and the resulting motion is seen to be periodic with the period *T*. In a similar manner the most general possible combination may be treated.

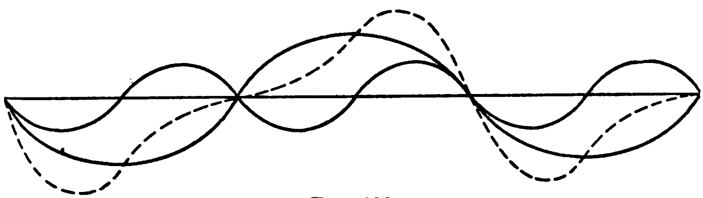


FIG. 108.

The exact nature of the vibrations of musical instruments, and the discussion of the individual properties of each instrument, will be given later, in Chapter V.

The essential characteristics of any periodic vibration are, then, (1) frequency; (2) amplitude; (3) complexity, or nature of component harmonic vibrations.

115. Measurement of the Period of Vibration. It is not difficult to measure the period of vibration of a vibrating body, or the reciprocal of this quantity, the frequency; and various methods are taught in laboratories. One method is to attach a bristle to the vibrating body, and to move under the bristle a piece of smoked glass at a known velocity. (E. g., fasten the glass to a pendulum and measure its period of vibration, or allow the glass to fall vertically downward toward the earth.) Then count the number of vibrations registered in a measured space on the glass, and calculate from the known velocity of the glass what portion of a second

has been taken in moving this distance; the ratio of this time to the number of vibrations counted is the period of the vibrating body. Another method is to compare the period of the body with that of a simple pendulum directly by means of "coincidences." Again, some instruments are so constructed as to register automatically the number of vibrations. The best instrument of this type is a so-called "siren." This consists of a circular disc which contains several series of openings arranged in concentric circles, and which can rotate in its own plane around an axis through its centre. These openings do not pass vertically through the disc, but each one slants tangentially slightly; so that, when a blast of air is blown through it, there is a pressure produced on the walls of the passage tending to make the disc revolve. If, then, a blast is forced up through a tube directed so as to strike a series of openings, there will be a succession of impulses given to the air above the disc; because, being set in motion by the blast, the disc will revolve, and so the openings will come in turn over the mouth of the tube conveying the air. The number of impulses per second given the air equals the number of openings multiplied by the number of revolutions per second of the disc. By altering the number of openings or the pressure of the air-blast, the number of vibrations can be changed; and it is perfectly simple to have the entire number of revolutions of the disc registered mechanically by a screw and dial. So the frequency of the vibration can be at once deduced. Another method will be described later (see Art. 126).

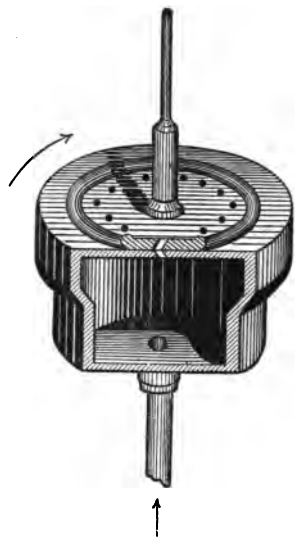


FIG. 109.

CHAPTER II

SOUND WAVES

As already stated, an interval of time elapses between the vibration of the sounding body and the perception of sound in our ears; and the presence of ordinary matter is necessary to carry the effect from the body to the ear. In this chapter the method of the propagation of this effect will be discussed.

116. Existence of Waves. As the sounding body vibrates, it gives a series of impulses to whatever matter is touching it; and so a disturbance is sent out into the surrounding medium. Thus, a tuning-fork vibrating in air or in water sends out waves; if it touches a solid like an iron pipe or a wooden table, it sends waves out in them also; because, by definition, a wave is simply the advance of a disturbance into a medium. See Articles 74 *et seq.*

117. Nature of Sound Waves. Since all sounding bodies are making vibrations, to and fro, they produce simply a "push and pull" in the surrounding medium. Therefore, the portions of the medium move backward and forward in the direction of the advance of the wave; that is, the waves which produce sound are *longitudinal* or "compressional."

Consider more in detail the simplest case, a tuning-fork vibrating in the air. Each time the prongs move out they produce a condensation of the air immediately in front; and owing to the elasticity of the air this condensation spreads out into the air. Then, as the prongs return, they produce a rarefaction, which also extends out into the air.

So, as the vibrations continue, there is a succession of condensations and rarefactions sent out. The individual particles of the air do not, of course, move onward with the wave, but vibrate backward and forward in the direction of the advance of the waves.

A very good illustration of this particular kind of wave is given by a model consisting of a series of heavy spheres suspended in line at regular intervals by long threads and separated by light spiral springs. If the first sphere is moved toward the right, it compresses the first spring; then the second sphere is moved; this compresses the second spring, etc. The compressional wave advances along the row of spheres with a velocity depending upon the mass of the spheres and the elasticity of the springs. Similarly, if the first sphere is moved toward the left, the spring is stretched; and so a rarefaction advances along the line, with the same velocity as the compressional wave did. If the first sphere is made to vibrate periodically, there will be a succession of condensations and rarefactions sent out. The individual spheres make periodic vibrations, as is evident, but do not change their positions of equilibrium.

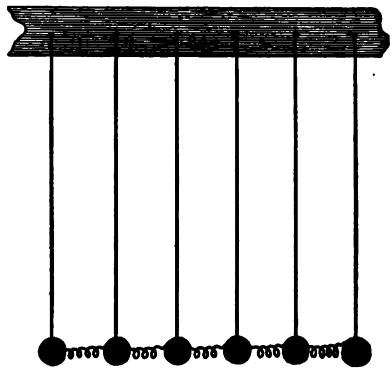


FIG. 110.

If the waves are in any other medium than air, they also consist in the advance of condensations and rarefactions.

118. Characteristics of the Waves. When a train of waves advances into any medium, it has certain fundamental characteristics. The wave-length, λ, is the distance from

condensation to condensation, or from any particle to the next particle in the direction of the advance of waves, which is in identically the same position and condition relative to its vibration. The "frequency" or "wave-number," n, is the number of condensations (or rarefactions) sent out in one second, or, what is the same thing, the number of condensations (or rarefactions) which pass by a fixed point in the medium during one second. The velocity, v, is the distance the waves advance in one second. Hence $v = n\lambda$. (This velocity has, of course, no connection with the speed of the individual particles of the medium; for the latter is changing each instant, and depends upon the intensity of the vibrations of the sounding body.) The "amplitude" is the length of the path through which each particle of the medium vibrates. The waves carry away energy from the sounding body; and the intensity of the waves (see Art. 74) is inversely proportional to the square of the distance from the source and directly proportional to the square of the amplitude. (This statement refers to the intensity of the waves, *not* to the intensity of the sound sensation.)

119. Velocity. As is evident, the velocity of the waves depends upon the elasticity and the inertia of the medium; and it may be proved that for any elastic wave the value of the velocity in an isotropic medium is

$$v = \sqrt{\frac{E}{\rho}},$$

where E is the particular coefficient of elasticity corresponding to the strain characteristic of the waves, and ρ is the density. It is seen, then, that the velocity is entirely independent of the length or the amplitude of the wave, since E and ρ are in general constants. Therefore, since $v = n\lambda$, if the frequency is small, the waves are long; and conversely. In sound waves, the strain is one of change of volume; and so E is the bulk-modulus (see Art. 81).

Further, the compressions and rarefactions are very rapid; and so no heat can enter or escape.

For a gas, then, the coefficient of elasticity is γp (see Art. 109), where p is the pressure and γ is a constant for any one gas. Thus

$$v = \sqrt{\gamma \frac{p}{\rho}} = \sqrt{\gamma R T},$$

because, as will be shown later (Heat, Art. 178) for a gas $p = R \rho T$, where R is a constant for any one gas, and $T = 273 + t^\circ$ on the centigrade scale of temperature. The velocity in a gas is thus independent of wave-length, of amplitude, and of density of the medium, but varies as the square root of $273 + t^\circ$ C. As will be shown later, v can be easily measured for any gas, so can R and t; consequently γ for that gas may be thus determined; and this in fact is the method generally adopted.

120. Composition of Waves. If the sounding body is making harmonic vibrations, it will send out what may be called a harmonic train of waves. These waves will be spherical; but, at some distance from the vibrating body, they may be regarded as practically plane. (See Art. 75.) Each particle of the medium is making harmonic

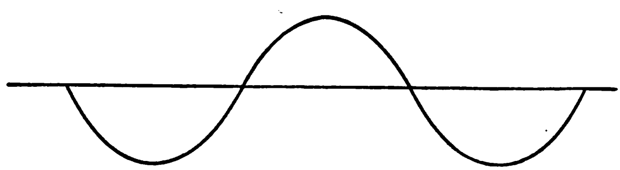


FIG. 111.

vibrations through a position of equilibrium; and so the motion is handed on. This may be represented graphically in this way: Draw a line in the direction of the advance of the waves; lay off on it points to represent the positions of equilibrium of various particles of the medium; erect

at these points perpendicular lines whose lengths equal the displacements of those particles at any instant from their positions of equilibrium; the locus of the extremities of these lines may be called the "wave-form." If the vibration is harmonic, the wave-form will be a sine-curve, as shown. In a sound wave, the particles are moving in the direction of the advance of the wave, not across the direction; but for the sake of clearness in the diagram the displacements are represented as if they were transverse.

At any instant later, the wave-form will be simply this first form moved forward in the direction of advance of the waves; and so the entire phenomenon of the train of waves may be considered graphically as the motion forward of this wave-form with the velocity of the waves.

If the sounding body is making periodic vibrations, which are not harmonic, these may be resolved, as expained in Article 114, into a series of harmonic vibrations of periods, T, $T/2$, $T/3$, etc.; and each of these partial harmonic vibrations may be considered as sending out a harmonic train of waves. So the resulting train of waves, due to any periodic vibration, may be regarded as the sum of a number of harmonic trains. The form of the waves will, of course, depend upon the amplitudes of these separate waves. Consequently, several trains of waves may have the same period and amplitude, but different "forms."

Similarly, if two or more sounding bodies are sending out waves into the same medium, the resultant displacement at any point will be the geometrical sum of the displacements which each wave separately would have produced, provided only that these individual displacements are small in comparison with the length of the waves. This is perfectly analogous to the combined action of two overlapping water-waves whose crests are not too high.

121. It is thus evident that there are three essential properties of a given train of waves: first, its wave-number

or frequency; second, its amplitude; third, its form. If the frequency is known, so is the wave-length, because in a given medium the velocity is the same for waves of all wave-lengths; and $T = n\,\lambda$. Therefore, for a given value of n, there is a definite value of λ in any one medium.

These three characteristics of a train of waves correspond perfectly to the three characteristics of a vibration: frequency, amplitude, and complexity.

CHAPTER III

SOUND SENSATION

122. Perception of Sound. These compressional waves, sent out by the vibrating source, spread through the surrounding matter; and, if a being with the sense of hearing is in the neighborhood, the waves will, in general, produce in his ear the sensation of sound. All compressional waves do not produce sounds; because it may be proved by experiment that, if the frequency is too small or too great, no sensation is perceived. These limits vary greatly for different individuals; but roughly it may be said that, if the frequency of the waves is between 20 and 40,000, sounds will be produced in the human ear. If the frequency of the vibrating body is less than twenty vibrations per second, it is doubtful if a true compressional wave is produced in a fluid; because, between consecutive vibrations, the fluid will have time to flow around the vibrating body, and will not be simply compressed by it. If the frequency is greater than forty thousand (in fact even if it is considerably less) the waves, although they may produce no sound in the ear, may be traced, and their laws studied, by means of what is known as a "sensitive flame." If any illuminating gas is allowed to escape into the air from a vessel which holds it under pressure, it will form a jet as it comes out; and, if this is lighted, and the pressure so regulated that the jet does just not flicker, its stability is exceedingly unstable. If a body, even some distance away, makes vibrations of a great frequency, the jet will collapse, owing to the effect of the very short waves

upon the surface which separates the jet from the air. The particularly sensitive point of the jet is near its base, where it emerges into the air.

FIG. 112.

123. The Ear. The waves enter the ear through the air, and necessarily set in vibration the "drum" of the ear, which closes the passage. This membrane is connected by a series of bones with the inner cavity of the ear, which contains a liquid substance. And it is in the inner ear, within the influence of this liquid, that the auditory nerves end. Consequently any vibrations of the ear-drum produce corresponding motions around the nerve-endings; but what the exact connection is between these two, i. e. the cause and effect, is at present unknown.

124. Musical Notes. The simplest vibration possible for a sounding body is, as stated above, a harmonic motion. It is a fact of experiment, too, that the simplest sound sensation known to us is that produced by an instrument which vibrates harmonically. Such a sound is called a "tone." If an instrument is vibrating any other way, it emits a train of periodic waves, which produces in the ear a more complex sensation. But it was stated that any such train of periodic waves may be considered as the sum of a number of harmonic waves; and it is also a fact of experiment that, when a train of periodic waves comes into the ear, the sound sensation is not a single one; but, on the contrary, several simple tones may be distinguished, and the sensation is thus complex. In fact it seems to be a general law that, no matter how many

trains of waves enter the ear, or how complicated they are, the complex sound heard may be mentally resolved into the simple tones corresponding to the harmonic waves composing the waves. This statement is sometimes called "Ohm's Law of Sound Sensation."

125. Characteristics of Musical Notes. If two musical notes are compared, the sensations differ in three respects. One may be shriller or "higher" than the other; one may be louder or more "intense" than the other; and finally, there is a marked difference depending upon the instrument which produces the sound.

126. Pitch. The "pitch" of a note is the name given to that property which marks its shrillness. And it is easily proved by experiment that, as the number of waves per second entering the ear increases, so does the pitch of the sound sensation grow higher. If the distance between the ear and instrument emitting the waves is not changing, the number of waves entering the ear equals the number of waves sent off by the instrument. Two vibrating bodies, then, under these conditions, which have the same frequency, produce sounds of the same pitch. That is, the pitch of the sound may be defined numerically as being equal to the frequency of the vibration or the frequency of the train of waves. A person with a musical education can easily distinguish between two notes whose pitches are extremely close; and so the frequency of any vibrating body may often be determined by comparing the pitch of the sound which it produces with that produced by some instrument whose frequency can be accurately measured and also varied. The pitch is slowly altered until its note and that of the first body have the same pitch, that is, are in "unison." Such an instrument, whose frequency can be varied continuously through a wide range, and can at the same time be accurately measured, is the "siren."

127. If, however, the vibrating body is approaching the ear or receding from it, the number of waves which enter

the ear in one second no longer equals the number emitted by the body. There is a greater or a less number of condensations or "crests" in a given space in the intervening medium than there would be if the body were not moving; that is, the wave-length is changed.

Let n = frequency of sounding body.
V = velocity of compressional waves.
v = velocity of body toward the ear.

Since V is the velocity of the waves, this means that in one second a crest which is at a distance V from the ear advances to the ear; and the number of crests, then, in this distance V enters the ear in one second. Owing, though, to the advance of the sounding body, the n waves or crests which it emits in one second are comprised in a space $V-v$; and so in a space V there are n' crests, where $n' : n = V : V-v$. Hence the pitch of the sound heard is

$$n' = \frac{nV}{V-v}.$$

Similarly, if the sounding body is receding from the ear with the velocity v, the pitch of the sound heard is

$$n' = \frac{nV}{V+v}.$$

This explains the change in pitch of the whistle of a locomotive when it approaches or recedes from an observer.

If the sounding body is at rest, and the person listening is approaching or receding, the case is somewhat different. Let n and V be as before, and let v be the velocity of the person away from the sounding body. There are n crests in a distance V; and so, as the person moves away from the source of the waves a distance v in one second, the

number of crests which overtake and pass him in one second is n', where

$$n' : n = V - v : V.$$

Hence the pitch of the sound is heard

$$n' = \frac{n(V - v)}{V}.$$

Similarly, if the person is approaching the sounding body with a velocity v, the pitch of the sound heard is

$$n' = \frac{n(V + v)}{V}.$$

This statement about the change in pitch of a sound due to the alteration in position of the sounding body and the hearer is called "Döppler's Principle;" and it may obviously be extended to any system of waves in any medium.

128. Intensity. The loudness or intensity of a sound may be shown to vary with the intensity of the waves; that is, it varies with the remoteness of the sounding body and with the amplitude of the waves. But one is not a measure of the other. In fact, as yet, no numerical value can be given the intensity of a sensation; and it is doubtful if in any case it would be exactly proportional to the energy received by the ear. All that can be said is that the intensity of any sound may be increased by increasing the intensity of the waves, and *vice versa*.

129. Quality. That difference which may exist between sounds of the same pitch and comparative intensity, and which enables one to distinguish between them, is said to be a difference in "quality." It has been shown by careful experiments to depend upon the "form" of the waves, that is, upon the partial vibrations which are superimposed upon the fundamental. Every instrument has particular partials which it produces; and they and their relative

intensities vary with different instruments, thus producing differences in the forms of the waves. Any vibration may be analyzed into its fundamental and upper partials by means of "resonators," which are instruments designed to pick out from any complex sound its component parts. The principle made use of is a most general mechanical one: If a train of waves of a certain frequency is passing through any medium in which is a body which, when it vibrates, has the same frequency as the wave, then this body will itself be set in vibration by the waves. The body will begin with a very minute motion; but the suc-

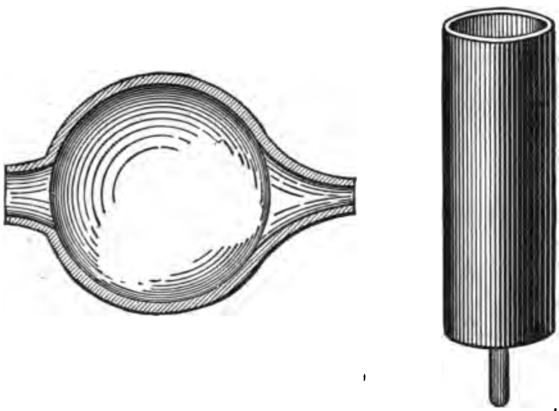


FIG. 113.

cessive impulses given by the waves are so timed as to further increase the motion; just as, by pulling at the proper instants, a man may set a heavy bell in vibration. An illustration of this principle is given when two tuning-forks of the same frequency are placed near each other, and one is set in vibration by means of some blow. If this one's motion is now stopped, it will be noticed that the other one is in vibration. The commonest form of resonator is a hollow sphere which has two small openings on opposite sides. A long cylinder, closed at one end, is

also often used. The air inside may be set in vibration by blowing over the mouth of the opening; and it has a definite frequency depending upon its size and shape. If, now, a vibration of this same pitch is given by a neighboring body, the air in the resonator will be set in "sympathetic" vibration; and so the sound heard will be strengthened. A vibration of any other frequency would not be thus reinforced. So, by having a series of resonators of different frequencies, any complex vibration may be resolved into its components; and it may thus be proved that differences in quality are due to differences in the number and intensity of the partials present. It should be remembered that if the fundamental of any note has a period T, its upper partials have the period $T/2$, $T/3$, etc.; that is, if the frequency of the fundamental is n ($= 1/T$), the frequencies of the partials are $2n$, $3n$, etc.

130. It is thus seen that, corresponding to the three physical properties of a train of waves, its wave-number, amplitude, and form, there are three characteristics of any musical note, — its pitch, intensity, and quality.

131. Combinations of Notes. If two or more trains of waves are passing through the same medium, the resulting displacement at any point is, of course, the geometrical sum of the individual displacements which each train of waves by itself would have produced. If two trains of waves so overlap that at some point the "crest" of one coincides with the "hollow" of the other, — that is, if the displacements due to the two waves are in opposite directions at that point,— there will be no motion at that point if the amplitudes of the two waves are equal. But if two crests coincide, the displacement will be abnormally great. This phenomenon of the overlapping of waves is sometimes called "interference;" and, when one wave neutralizes another at any point, the interference is said to be complete. There are corresponding properties of sensation.

132. Beats. If two sounding bodies, e. g. two tuning-

forks, of different frequencies are vibrating near each other, each sends out a train of waves which reaches the ear of an observer near by. But the wave-form of the combined motion contains places of no motion if the amplitudes are the same, and places of abnormal motion, as is shown in the diagram, which represents the combination of two harmonic trains of waves whose frequencies are in the ratio of two to three. In general, since one train of waves has more "crests"

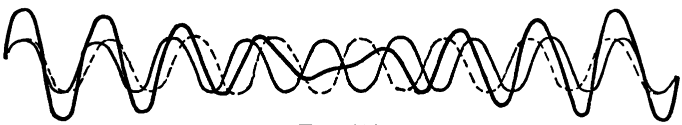


FIG. 114.

in a given space than the other, they will completely interfere at certain points, the number of which in a distance equal to the velocity of the waves, equals the difference of the two frequencies. As this wave-form reaches the ear, the membrane closing it does not vibrate regularly; at certain intervals it is violently stimulated, while at times in between it is not moved at all. This necessarily produces in the ear a sensation analogous to that caused in the eye by an intermittent series of flashes of light; and both these sensations are disagreeable. The degree of the unpleasantness depends upon the number of stimuli given the ear or eye in a given time, and also upon the frequencies of the waves. In sound these unpleasant effects are called "beats;" and, if they are produced by the interference of two waves whose frequencies are n and n', there will be $n' - n$ beats per second, and the unpleasantness will vary with $n' - n$ and also n. That is, a certain number of beats per second may be more disagreeable at one pitch than another. If the number of beats exceeds forty or fifty in a second, it becomes difficult to detect them; and so the most disagreeable effect is produced when the pitches of the two sounds are nearly the same. If the pitches are identically the

same, there are, of course, no beats; and their absence may serve to demonstrate the unison of two notes.

133. Differential and Summational Notes. When two vibrating bodies of frequencies n and n' are set in motion violently, other sounds than those of pitches n and n' are heard, if the same volume of air is agitated by the two vibrations. Among these so-called "combination-notes" are those of pitch $n + n'$ and $n - n'$, which are called respectively "summational" and "differential" notes. Other notes are also heard under different conditions.

CHAPTER IV

Reflection and Refraction

134. Introduction. Waves of all kinds, moving in any medium, suffer certain changes when they reach the bounding surface which separates that medium from another. When the waves reach the boundary, they produce "reflected" waves back into the first medium, and also cause waves to pass into the second medium. Those waves in which these facts are most easily observed are light-waves, where the paths of the waves may be seen; and their properties will be studied later. But every one is acquainted with a few of the fundamental laws of light: the angle of incidence equals the angle of reflection; when light-waves pass from one medium into another, the direction of the waves is changed, i. e. suffers "refraction;" concave mirrors and convex lenses can focus the light-waves; there are certain so-called interference and diffraction phenomena. All these properties are common also to those waves in ordinary matter which can produce sound, i. e. to all compressional waves.

That sound waves can be reflected is made evident by the existence of echoes and "whispering-galleries." Concave mirrors can also easily focus sound waves; and a "sounding-board" is an illustration of this.

Lenses can also be made of suitable size, so that the waves from a vibrating body placed at A will all converge, after passing through the lens, at some point, B. A and B are known as "conjugate foci." (See Fig. 115).

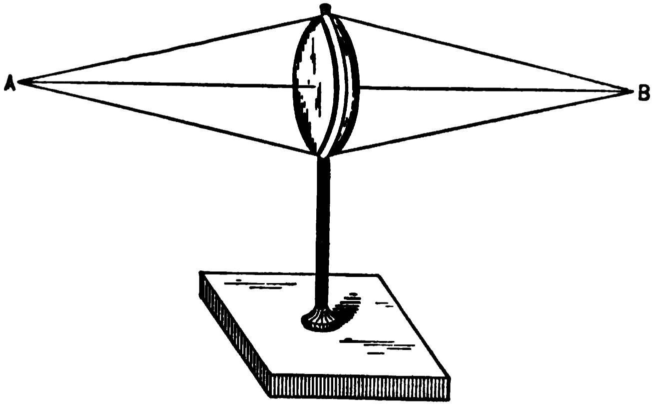


FIG. 115.

135. Reflection. The nature of the reflection of a train of waves of any kind depends on the differences between the two mediums at whose boundary the reflection takes place. This can best be explained by some mechanical illustrations. Consider two series of spheres arranged in

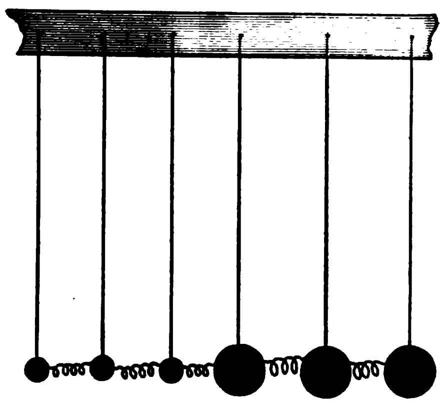


FIG. 116.

line and separated by spiral springs. Let the masses of the spheres and the elasticity of the springs be such that the velocity of a compressional wave is different in the

two series. Call one the "fast" series; the other, the "slow."

Now, if the extreme sphere of the slow series is pushed in so as to compress the spring attached to it, a compression is sent along the row; but, when the wave reaches the end of this series, the velocity of the compression in the fast series is such that the last sphere of the slow series does not meet with enough opposition; and so it swings further than it should, and thus stretches the spring between it and its next neighbor in the slow series. Consequently a rarefaction is sent back as a reflected wave along the slow series of spheres. But, if a compression is sent along the fast series, when the last sphere is reached, it compresses the spring between it and the first sphere of the slow series; and, since the velocity in this series is less than along the other, the spring does not move fast enough, and so becomes unduly compressed. Consequently by its reaction on the fast series a compression is sent back as the reflected wave. Similarly, if a rarefaction is sent along the slow series, a condensation is reflected; and if it is sent along the fast series, a rarefaction is reflected. So, if a succession of condensations and rarefactions is sent along either series, a train of waves is reflected.

A special case of a series of spheres, whose velocity of transmission is slow, is one where there is no transmission at all, e. g. the last sphere of one series is fastened to a rigid wall. In this case the reflected wave is identical with the incident, only going in the opposite direction. A special case of a series of spheres whose velocity of transmission is very rapid, i. e. where little opposition is offered to the incident wave, is when there is no second series at all; that is, there is a single series of spheres, and the last sphere is not connected with anything except the spring which gives it motion. In this case the reflected train of waves differs from the incident in that a condensation is replaced by a rarefaction, and *vice versa*.

Transverse Waves in Cords. Another illustration of the reflection of waves is given when transverse waves are sent along a rope or stretched string. Consider the two cases, (1) reflection from a fixed point, i. e. the further end is rigidly fastened; (2) reflection from a "free" end, i. e. the further end is not fastened to anything.

136. 1. **Reflection from a Fixed Point.** Let a transverse harmonic wave be sent along a cord whose further end is fastened to a wall or any fixed support. At the instant it reaches the fixed end, it may be represented as in the figure, where the dotted line shows the original position of the cord. *B* is the fixed point, and *A* is the end which is moved transversely with a harmonic vibration. The instant this wave reaches *B*, a similar wave starts back in

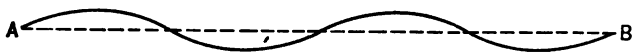


FIG. 117.

the opposite direction along the cord, so that, at later instants, as the end *A* is vibrated, thus sending out waves in the cord, there are two trains of waves moving in opposite directions, — one direct, the other reflected. These two trains have the same frequency and velocity; and their positions are shown in the following diagrams, at instants of time differing by one eighth of the period of the wave. The faint continuous line represents the direct wave; the dotted line, the reflected wave; the heavy line, the resulting motion produced by the sum of the displacements due to each of the two waves separately.

137. It is to be noticed that at regular intervals there are points of the cord which never move, the two waves neutralizing each other. These points are called "nodes;" and the distance between consecutive nodes is seen at once to be one half a wave-length of either train of waves. Also, in between the nodes, it is noticed how the portions of the cord vibrate as a whole across the original direction.

The middle point of each such vibrating portion is called a "loop," and the distance between two consecutive loops is also one half a wave-length. There will thus be produced in the cord, as a consequence of the superposition of these two waves, a transverse *vibration*. This is called a "stationary" vibration, because, although at any instant the cord has the appearance of a train of waves, this apparent wave does not advance; the crests remain at the same points, die down, and then hollows take their places, while the hollows rise and become crests; and so on.

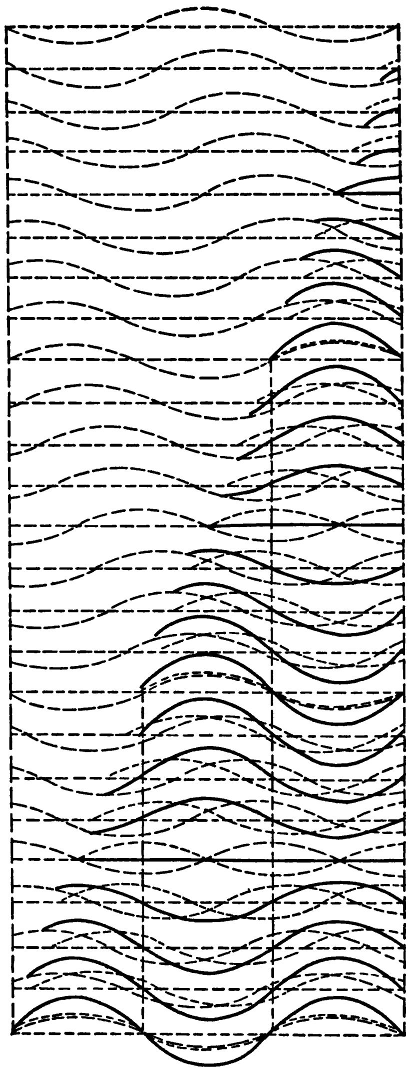


FIG. 118.

If a cord is thus set in stationary vibration by the hand, it is noticed at once that the end of the cord in the hand is approximately a node. It does move some, because the energy from the hand is here given the cord; but its motion is not nearly so great as at a loop.

138. Thus, in a cord fastened at one end, and set in stationary vibrations by a hand at the other, there are nodes at each end. (So it is also if a cord is fastened at both ends and set in vibration by some blow.) Consequently, there must be a whole number of vibrating portions or segments in the cord. If there are $N+1$ nodes, counting the two ends, there are N such vibrating segments. Let v be the velocity of the two trains of waves in the cord; n, the frequency; λ, the wave-length; L, the length of the string. Then

$$v = n\lambda; \quad L = N\frac{\lambda}{2},$$

since the distance between two nodes is $\frac{\lambda}{2}$. Consequently,

$$n = \frac{Nv}{2L}.$$

Therefore, since N can equal 1, 2, 3, 4, etc., there are certain definite values of n which can produce stationary vibrations in a given cord. To make N greater, — that is, to increase the number of vibrating segments, — it is necessary to increase n, the number of vibrations per second given the cord.

If the tension in the cord is increased by some longitudinal force, the velocity, v, will be increased. Consequently, if n is not changed, and if the tension in the cord is increased enough, N must decrease by 1. That is, if a cord is vibrating with 4 segments, it will change into 3 segments, if the tension is increased sufficiently.

All these deductions can be easily verified by direct experiments.

139. 2. **Reflection from a Free End.** Let a transverse harmonic train of waves be sent along a cord whose farther end is free. The reflected waves will have the same frequency and velocity, and they will be identical with the reflected waves from a fixed point, except that a crest takes

the place of a hollow and *vice versa.* The following drawings represent the direct train of waves, the reflected train, and the resulting motion, at intervals of time equal to one eighth of a period.

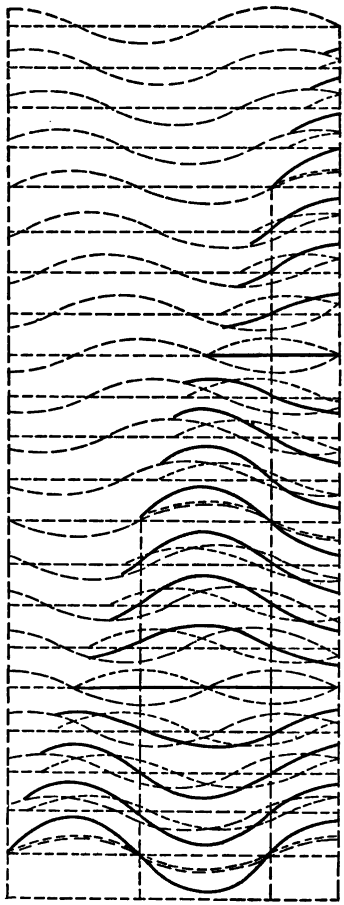


FIG. 119.

It is seen that here also there are nodes and loops, i. e. a stationary vibration; and the distance between two nodes or between two loops is equal to one half the wave-length of either the direct or the reflected waves. There is a loop, of course, at the free end of the cord.

140. Longitudinal Waves. Just as in the ball and spring model discussed in Article 135, so *longitudinal* waves may be sent along a stretched cord, a solid rod, a column of air. These waves are, of course, compressional ones, and so move with the "velocity of sound" in those substances. The waves will be reflected from the ends of the medium; and thus it will be set in stationary vibration. If the end is fixed, there will be a node there; if it is free, there will be a loop. The number of vibrating segments depends upon the velocity of the waves and the number of vibrations per second.

CHAPTER V

VIBRATING BODIES

141. Nearly every musical instrument when it is producing the compressional waves commonly called sound-waves is making stationary vibrations, the nature of which has been discussed in the previous chapter. The fundamental property of stationary vibrations is the existence of nodes and loops which may be regarded as produced by the superposition of two identical trains of waves in opposite directions; and it is a general law that the wave-length of both the direct and reflected waves is twice the distance between two nodes or between two loops. As explained also, any body such as a cord or a column of air can vibrate in many ways, with one segment or two or three, etc.; and so in general, unless special precautions are taken, a body set in vibration at random will have a motion which is compounded of several of these different simple modes of vibration. Some special cases will now be considered.

142. Transverse Vibrations of a Uniform Stretched Cord. When a cord fastened at both ends vibrates transversely, it itself does not affect greatly the surrounding air, owing to its small cross-section. But the supports to which the ends are fastened are never quite rigid; and so they also vibrate, and in turn produce vibrations in the board to which they are joined. This board may be made broad; and consequently its vibrations will produce a marked effect on the surrounding air. This explains the action of a violin or a piano, where the wooden portions are much more important than the strings.

There are many ways in which a string thus stretched can be set in transverse vibrations, — it may be plucked one side, it may be rubbed with a violin bow, or it may be struck a blow. There are also many possible modes of vibration. The simplest one is when the only nodes are at the two ends. In this case, if L is the length of the string, n the frequency, v the velocity of *transverse* waves in the string, the wave-length, v/n, equals $2L$. That is,

$$n = v/2L.$$

This vibration is called the "fundamental." The next simplest mode of vibration is when the string is divided

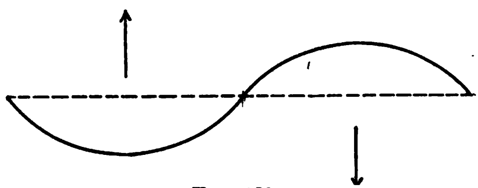


FIG. 120.

into two vibrating segments, there being a node at the middle point. If n_1 is the frequency of the vibration,

$$v/n_1 = L,$$

or

$$n_1 = v/L.$$

This vibration is called the "first partial;" and its frequency is just twice that of the fundamental. The next

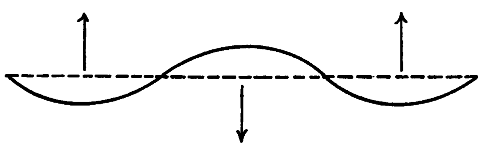


FIG. 121.

simplest case is when there are three vibrating segments. If n_2 is the frequency, it is obvious that

$$v/n_2 = 2L/3,$$

or

$$n_2 = 3v/2L.$$

This vibration is called the "second partial;" and its frequency is three times that of the fundamental. In a similar manner, it is easily seen that there are other possible partials, whose frequencies are four, five, etc. times that of the fundamental.

Corresponding to each of these partial vibrations there is a definite sound-sensation, whose pitch is given by the frequency of the vibration; and, if the string is set in motion at random, several of these sounds may be heard. Of course the point of the string, at which is struck the blow producing the vibration, is always a loop in the resulting vibration; and so any partial vibration which would make it necessary for that point to be a node cannot occur, and the corresponding sound is absent. Thus, if a string is set in vibration by a blow at its middle point, the first, third, fifth, etc. partials cannot arise. Since the quality of any sound depends upon the number and intensity of the partials present with the fundamental, the sound produced by the transverse vibrations of a stretched string depends largely upon the point where it is struck or bowed.

It is proved by theory and verified by experiment that the velocity of a *transverse* wave in a thin, uniform stretched string is

$$v = \sqrt{\frac{P}{\sigma\rho}},$$

where P is the tension in the string, that is, the stretching force; σ is the area of the cross-section of the string; ρ is the density of the string. Consequently the pitch of the fundamental, $n = v / 2L$, is

$$n = \frac{1}{2L}\sqrt{\frac{P}{\sigma\rho}};$$

and so, as the tension is increased, the pitch rises.

143. Longitudinal Vibrations in a Stretched Cord or Wire. This mode of vibration may be produced by rubbing the cord in the direction of its length with a damp cloth. The waves thus excited, which by their combination produce the vibration, are compressional; and so the velocity of the waves is what is called the velocity of sound in the material of which the cord or wire is made.

The simplest case, as before, is that where the only nodes are at the two ends. If L is the length of the cord or wire, n the frequency, V the *velocity of sound* in the given material,

$$V/n = 2L,$$

and so

$$n = V/2L.$$

Since n and L can both be measured, this gives a method of determining V, the velocity of sound in the given substance. Other cases of vibration, where there are nodes along the cord or wire, are obtained with difficulty.

It may be proved that, in this case,

$$V = \sqrt{\frac{E}{\rho}},$$

where E is Young's modulus of elasticity (see Art. 83), and ρ is the density of the cord or wire. Then the frequency of the fundamental vibration is

$$n = \frac{1}{2L}\sqrt{\frac{E}{\rho}},$$

or

$$E = 4L^2 n^2 \rho.$$

And since L, n, and ρ are all easily measured for any given cord or wire, this gives a simple method for the determination of Young's modulus. It should be noticed that the frequency of the vibration, for a definite length, does not depend upon the cross-section of the cord or wire, nor upon the tension in it, provided that this is not great enough to influence Young's modulus or the density.

144. Longitudinal Vibrations of Rods. This mode of vibration may be produced by rubbing the rod; and it is exactly similar to the preceding case, with this exception: a stretched cord or wire is fastened at its two ends, while a rod is not in general.

The simplest longitudinal vibration of a rod is when it is clamped at its middle point. There is then a *loop* at each end; and the wave-length

$$V/n = 2L,$$

or

$$n = V/2L.$$

Here, also, n and L may be measured; and so this gives a method for the measurement of the velocity of sound in the material of which the rod is made.

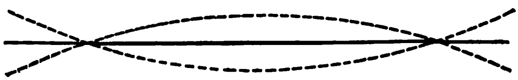


FIG. 122.

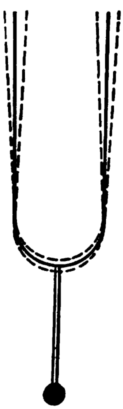


FIG. 123.

145. Transverse Vibrations of a Rod. The rod may be clamped at some point, and then struck a blow at one end perpendicular to its length. One of the commonest modes of such a vibration is when one end is clamped, and the other vibrates freely. In this case there is a loop at the latter end and a node at the former. Another mode of vibration of a rod is when there are two nodes at equal distances from the ends. An illustration of this is an ordinary tuning-fork, which is simply a rod so bent as to form a U, and which has an arm attached at the bottom of its curved portion. When a tuning-fork vibrates, its two ends are loops; and the attached arm is also in violent motion, as may be proved by resting it on a table or box; for the latter is immediately set in vibration.

146. Vibrations of Columns of a Gas. The only vibrations possible in a column of gas are, of course, longitudinal ones, because a gas has no rigidity. These vibrations are extremely common in musical instruments, e. g. organ-pipes, horns, flutes, etc.; and they are easily produced. One way is simply to blow over the opening of a hollow tube; another is to hold some vibrating instrument, like a tuning-fork, over the opening. In an ordinary organ-pipe air is driven in through a narrow passage, and, striking the lip at the bottom of the pipe, produces a periodic motion of the air near there, which sets in vibration the column of air in the pipe. There is always a loop at the lip of the organ-pipe, owing to the violent motion near it. Organ-pipes are of two kinds, — closed and open.

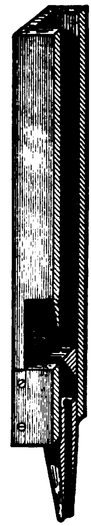


FIG. 124.

147. 1. Closed Pipes. The end of the pipe being closed, there is, of course, a node at that point.

The simplest mode, then, of vibration is when there is a node at the closed end and a loop at the lip, there being

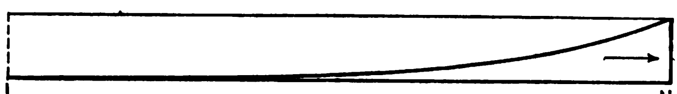


FIG. 125.

neither node nor loop in between. The wave-length is *four* times the distance from a node to a loop. So, if L is the length of the pipe, V the velocity of sound waves in the gas enclosed, and n the frequency,

$$V/n = 4L,$$

or

$$n = V/4L.$$

This is the fundamental vibration.

The next simplest mode of vibration is when there are a node and a loop in between the two ends. (It is, of course, impossible to have two nodes without there being a loop

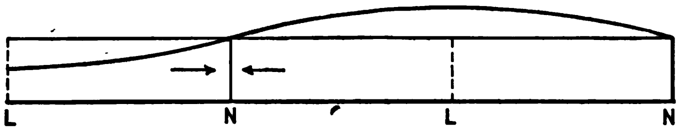


FIG. 126.

between them.) In this case the distance between two nodes is $2\,L\,/\,3$. Hence, if n_3 is the frequency,

$$V\,/\,n_3 = 4\,L\,/\,3,$$

or

$$n_3 = 3\,V\,/\,4\,L.$$

This is evidently the *second* partial, the first one not occurring.

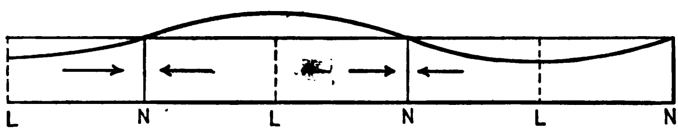


FIG. 127.

The next simplest mode of vibration is when there are two nodes and two loops in between the ends. If n_4 is the frequency,

$$V\,/\,n_4 = 4\,L\,/\,5,$$

or

$$n_4 = 5\,V\,/\,4\,L.$$

This is the *fourth* partial, the third not occurring.

So the other modes of vibration may be discussed. It is evident that in a closed organ-pipe only the fundamental and the alternate partial vibrations can occur.

148. 2. **Open Pipes.** In an open pipe there is a loop at each end. (In reality, owing to the inertia of the air, the loops do not come exactly at the ends, but at a slight dis-

tance beyond them, depending largely upon the radius of the pipe and the length of the waves.)

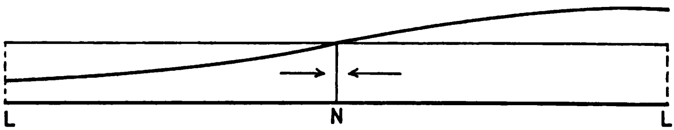


FIG. 128.

The simplest mode of vibration of an open pipe is when there is a loop at each end and one node in between. If n is the frequency,

$$V / n = 2 L,$$

or

$$n = V / 2 L.$$

That is, the lowest frequency of an open pipe is twice that of a closed pipe.

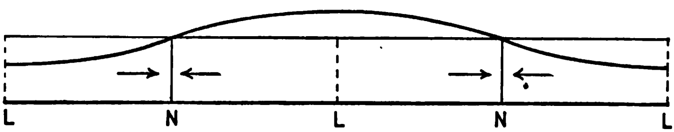


FIG. 129

The next simplest mode of vibration is when there are two nodes and one loop in between the ends. If n_1 is the frequency,

$$V / n_1 = L,$$

or

$$n_1 = V / L.$$

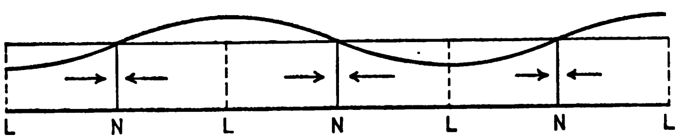


FIG. 130.

The next simplest mode of vibration gives

$$n_2 = 2 V / L,$$

and so on.

It is evident that in open organ-pipes the fundamental and all the partials may occur.

The quality of the sound heard as the result of the vibrations of the columns of gas in any organ-pipe depends upon the nature of the partials present. When an opening is made in the side of a pipe, as in a horn or a flute, a loop must occur at that point; and so the pitch and quality of the resulting sound are changed. By changing the motion of one's lips at the mouthpiece of a horn, different vibrations may be produced.

Further, since in a closed organ-pipe the fundamental vibration has a frequency $n = V/4\,L$, and since n and L can both be measured, V, the velocity of sound in the gas filling the pipe, may be easily determined. It is found, as might be expected, that the velocity in the pipe depends somewhat upon the size of the pipe and also upon the wave-length.

149. Vibrations of Plates and Membranes. A plate or a stretched membrane may easily be set in transverse vibrations by rubbing a violin bow across its edge. In general there will be certain *lines* in the plate or membrane, where there is no motion. These lines are called nodal lines; and their position may be found by having light sand or powder sprinkled over the surface. The sand will move away from the places of greatest motion and be heaped up along the nodal lines. As might be expected, the position and form of the nodal lines depend upon where the plate or membrane is held, and where and how it is rubbed.

150. Vibrations of Bells. A bell-shaped vessel, like a bowl of any kind, may have two kinds of vibrations, transverse and longitudinal. The transverse ones may be produced by striking the bowl a blow or by rubbing across its edge. The longitudinal ones may be produced by rubbing along its edge. As a rule, most actual bells, especially large ones, are not uniform, but have what may be considered as thicker places at certain points. If this thick place hap-

pens to be at a node, there will be a certain definite frequency; while, if it comes at a loop, the frequency will be slightly different. So, if the bell is struck at random, both these vibrations will occur, and beats (see Art. 132) will be heard.

151. The Human Voice. The vibrations of the lips, the tongue, and the so-called "vocal cords," or "larynx," produce the human voice. The pitch depends upon the frequency of the vibration. Various partials always occur; and these may be strengthened or modified by holding the lips and tongue in certain positions. The so-called "vowel" sounds are due to the simple vibration of the larynx and the modification of the partials by the shape given the mouth. The "consonant" sounds are due to the vibrations of the lips and tongue in different positions.

152. It should be noted that in all the cases and methods of vibration, which have been discussed, the wave-length of the sound waves varies directly as the *length* of the vibrating segment which produces the waves. This is a general property of waves: the wave-length varies directly as the *linear* dimension of the vibrating body.

CHAPTER VI

VELOCITY OF SOUND

WHAT is meant by the "velocity of sound" is the velocity of compressional waves in a material medium of a given kind and under given conditions.

153. Direct Method. This velocity may be determined without much difficulty by a direct measurement if the expanse of the medium is great enough. Thus a pistol may be fired in the air, or a bell rung under water; and the interval of time which elapses between this instant and that when the corresponding sound is heard at a measured distance away may be noted. The ratio of this distance to the interval of time gives the velocity of sound in air or water under the existing conditions.

154. Indirect Method. Another method is to apply the formula stated in Article 119,

$$V = \sqrt{\frac{E}{\rho}},$$

where E is the bulk modulus and ρ is the density, Thus, if E and ρ are known (and they can in general be determined), V can be at once calculated. For a gas, as proved in Article 119,

$$V = \sqrt{\gamma R T};$$

and, if V can be measured by any means for the gas, this formula may serve for the calculation of any of the other quantities. It should be noted that $T = 273 + t°$ centigrade; and so, as the temperature of a gas increases, the velocity of sound in it does also.

155. Velocity in a Column of Gas. In the last chapter other methods for the measurement of the velocity of sound in any medium have been explained. In particular, to measure the velocity of sound in any gas, enclose a column of the gas in a pipe open at one end, vibrate a tuning-fork of known frequency over this end, and alter the length of the pipe by means of a movable piston until the column of gas is vibrating in unison with the fork. Let n be the frequency of the vibration, L the length of the column of gas, and V the velocity of sound in the gas; then, if the vibration of the column of gas is its fundamental one,

$$4L = V/n,$$

or

$$V = 4nL.$$

An upper partial vibration of the column of gas might have been used; and, by suitable methods, all error due to the displacement of the loop from the open end of the pipe may be obviated.

If the velocity of sound in any one gas is known, it may be determined for any other gas in a most simple manner. Fill two organ-pipes with the two gases, and set the columns in vibration. Then the length of one can be altered until the pitches of the two sounds are identical. If the pipes are closed at one end and are emitting their fundamentals, and if n is the frequency common to both pipes, V_1, and V_2 the two velocities, L_1 and L_2 the lengths of the two columns,

$$n = \frac{V_1}{4L_1} = \frac{V_2}{4L_2},$$

and hence

$$V_1 : V_2 = L_1 : L_2.$$

So the unknown velocity may be at once calculated.

As noted before, it will be found that the velocity changes with the temperature, and that it depends slightly upon the relative dimensions of the wave-length and the organ-pipe.

156. Velocity in a Column of a Liquid. Stationary waves may also be produced in a closed pipe filled with a liquid, by causing the piston closing one end to vibrate harmonically. If n is the frequency, L the distance between two nodes, and V the velocity of sound in the liquid,

$$2L = V/n,$$

or

$$V = 2nL.$$

It is not difficult to determine accurately the position of the nodes; because, if any light powder is suspended in the liquid, it will arrange itself in the tube so that the nodes and loops may be easily distinguished.

157. Velocity in a Solid Wire or Rod. The velocity of sound in a solid wire or rod may be determined by simply setting it in *longitudinal* vibration. If the wire has a length L, and if it is stretched between two fixed clamps, then

$$2L = V/n,$$

and

$$V = 2nL.$$

The velocity of sound in two different wires may be also thus compared, by altering the length of one until its frequency equals that of the other. If V_1 and V_2 are the two velocities, L_1 and L_2 the two lengths,

$$V_1 : V_2 = L_1 : L_2.$$

If a rod is clamped at its middle point, and set in longitudinal vibrations, there are loops at the two ends; so, if L is the length of the rod,

$$2L = V/n,$$

or

$$V = 2nL.$$

158. Kundt's Method. Another indirect method of determining the velocity of sound in any medium — solid, liquid, or gas — is known as Kundt's method, from the name

of the physicist who devised it. Its principle is to compare the unknown velocity with one that is known. The apparatus consists of two parts: a glass tube, about 100 cm. long and of 2 or 3 cm. diameter, which is closed at one end by a tight-fitting piston, A, and at the other by a loosely fitting piston, B; and a solid rod, perhaps 100 cm. long, which is firmly clamped at its middle point, coaxial with the glass tube, and which carries on one end the piston B. The solid rod is set into longitudinal vibrations

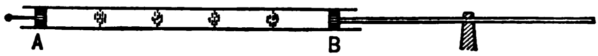


FIG. 131.

by being rubbed lengthwise with a dampened cloth; and the vibrations of the piston B set into vibration the gas enclosed in the glass tube. These vibrations in the gas will not become stationary unless the length of the column of gas is such as to equal a whole number of vibrating segments. Some light powder is placed in the tube, and, if the vibrations are not stationary, the powder will not take any permanent distribution. But, by moving the piston A, the length of the column of gas may be so changed that the vibrations will become stationary. This will be evident by the arrangement of the powder in the tube, and the distance between two nodes may be accurately measured. (It is to be noted that the point B is a node for the gas, although it is a loop for the rod; just as, when a rope is set in vibration by the motion of one's hand, the end in the hand is a node.) If L_1 is the length of the rod, V_1 the velocity of sound in the rod, L_2 the average distance between two nodes in the column of gas, V_2 the velocity of sound in the gas, n the frequency of the vibration, which is the same for both rod and gas,

$$2 L_1 = V_1 / n; \; 2 L_2 = V_2 / n,$$

or

$$V_1 : V_2 = L_1 : L_2.$$

So, if either V_1 or V_2 is known, the other may be at once calculated from the measurements L_1 and L_2. Thus, assuming that the velocity of sound in air is known, that of any solid may be at once determined by making a rod out of it, and using it as described. The velocity in any other gas or in any liquid may also be determined by first finding the velocity in any solid rod and then replacing the air in the tube by the other gas or the liquid.

This method is not so accurate as is desirable; but it is by far the easiest and simplest one available for general use.

TABLE II

Velocity of Sound

Air	0°	33,250 cm. per sec.
Hydrogen	0°	128,600 " "
Illuminating Gas	0°	49,040 " "
Oxygen	0°	31,720 " "
Alcohol (absolute)	8°.4	126,400 " "
Petroleum	7°.4	139,500 " "
Water	4°	140,000 " "
Brass		361,700 " "
Copper		397,000 " "
Glass		506,000 " "
Iron		509,300 " "
Paraffin	16°	130,400 " "

CHAPTER VII

HARMONY AND MUSIC

159. Musical Sounds. As already stated in Article 124, the simplest musical note that the ear can recognize is that produced by a harmonic vibration. It has also been known for at least two thousand years that there were certain combinations of sounds which were pleasant to hear. One such combination is called the "octave," which is due to the simultaneous production of two vibrations of suitable frequencies. Another combination consists of three vibrations, and is called the "major triad."

160. Numerical Relations. It was a matter of great interest to measure the frequencies of these vibrations which compose the pleasing combinations known for so many years. It was found that whenever a vibration and another of *twice* the frequency took place together, an octave was heard. Thus, if two tuning-forks of frequencies n and $2\,n$ are sounded simultaneously, an octave is heard; and the second vibration is often called the octave of the first.

Again, it was found that when three vibrations whose frequencies were n_1, n_2, n_3, were sounded together, a major triad was heard, only if

$$n_1 : n_2 : n_3 = 4 : 5 : 6.$$

Thus three tuning-forks whose frequencies are 200, 250, 300, will cause a major triad.

161. Harmony and Discord. The reason *why* these simple combinations of vibrations produce a pleasing sensation remained unknown until it was given by Von Helmholtz.

As explained in previous chapters, any vibration is, as a rule, always accompanied by partial vibrations whose frequencies in all simple cases are twice, three times, etc., that of the fundamental vibration itself; it has also been shown that beats are very unpleasant to the ear, and that the degree of the disagreeable sensation depends upon the number of beats per second and the pitch of the individual sounds. Thus, a vibration whose frequency is 200 is accompanied by partials whose frequencies are 400, 600, 800, etc. Consequently, if two vibrations, of frequencies 200, and 400, are produced simultaneously on different instruments, the following frequencies occur:

200, 400, 600, etc.
400, 800, 1200, etc.

and there are therefore no unpleasant beats between any of the vibrations. So it is seen why an octave, such as 200 and 400 form, is in harmony.

But if vibrations of frequencies 205 and 400 are sounded together, the following frequencies occur:

205, 410, 615, 820, etc.
400, 800, 1200, etc.

And now there are 10 beats per second between the two vibrations 410 and 400, also 20 per second between 820 and 800, etc.; and so this combination is unpleasant, and is said to be dissonant.

Again, if three vibrations forming a major triad are sounded together, e. g. 200, 250, 300, the following frequencies occur:

200, 400, 600, 800, etc.
250, 500, 750, 1000, etc.
300, 600, 900, 1200, etc.

No beats arise; and so the harmony of the combination is explained. If one of the fundamental vibrations was 205

instead of 200, there would be beats; and the resulting sound would be dissonant.

Further, Von Helmholtz proved that in every discord there were beats, generally between the partial vibrations; while in harmony the beats are almost entirely absent.

162. Musical Scales. The pitches of the sounds which compose an octave or a major triad are too far apart to allow any complicated music to be composed entirely of them; and so the attempt has been made at various times to introduce other sounds into music, and to make thus what is called a musical "scale." The simplest scale proposed is one built up of major triads and octaves; in the interval of an octave seven sounds are taken, whose pitches are proportional to the first seven of the following series of numbers:

$$24, 27, 30, 32, 36, 40, 45, 48, 54, 60, \text{etc.}$$

This series of numbers may be continued, as shown, by taking the octaves of the original series of seven.

It is seen that
$$\begin{aligned} 24 : 30 : 36 &= 4 : 5 : 6, \\ 36 : 45 : 2 \times 27 &= 4 : 5 : 6, \\ 32 : 40 : 48 &= 4 : 5 : 6, \end{aligned}$$

so that the scale is, as it were, composed of major triads; and it is called the "diatonic" scale. Of course any pitch may be taken as the starting-point, and from it the pitches of the other notes may be calculated. Thus it has been proposed to take as the standard a note whose pitch is 256; then the pitches of the other notes in the scale are found by multiplying each of the numbers 24, 27, 30, etc., by 256 / 24 or 32 / 3. In a similar manner the scale between 256 and the octave below it, 128, can be built up; and thus starting from any arbitrary pitch the scale both below and above it may be constructed. The standard pitch which is accepted in most countries to-day is 261, although this recognition is by no means universal.

The "interval" between two notes is the ratio of their pitches; and it is at once seen that the intervals in the diatonic scale are most uneven. The successive intervals are 9 / 8, 10 / 9, 16 / 15, 9 / 8, 10 / 9, 9 / 8, 16 / 15, etc. An interval 9 / 8 or 10 /9 is called a "tone;" that of 16 / 15, a "semi-tone." An attempt was made to equalize these intervals by introducing into each interval of a tone another note, which was called the "sharp" of the note before it, or the "flat" of the following note, and thus five notes were added.

But still the intervals were unequal, and so a radical change was made. There were introduced into the interval of an octave twelve notes, the intervals between successive ones being the same. Thus, if the standard note has the pitch n, and if the interval common to all the notes is a, the scale will be . . ., n, a^n, $a^2 n$, $a^3 n$, . . . $a^{11}n$, $a^{12} n$, . . . etc. But since there are twelve notes in an octave, $a^{12} n$ must equal $2\,n$, or $a = \sqrt[12]{2}$. This is called the "tempered" or the "chromatic" scale. In this scale the notes no longer form exact major triads; and so there is a faint discord in all music played on it; but our ears are so accustomed to it that it is rarely noticed.

BOOK III

HEAT

BOOK III

HEAT

INTRODUCTION

EVERY one is familiar with the sensations of heat and cold; and it is well known that the causes which produce these sensations can also produce other effects. The study of the nature of these so-called "heat-effects" and of the laws according to which they are produced forms what is known in Physics as the subject of "Heat."

163. It was noted, in speaking of Work and Energy (Art. 67), that, whenever work is done against friction, heat-effects are produced. Ice may be melted by friction; water may be heated until it boils by friction; a piece of wood may be set on fire by friction, if suitably applied. The amount of the heat-effect may be measured by the amount of the ice melted, the amount of the water boiled, etc.; and it is found, as the result of most careful experiments, that the amount of heat-effect depends only upon the amount of work done against friction.

Of course heat-effects may be produced in other ways than by friction. If a gas is compressed, its temperature is raised; and this is a well-known heat-effect. But here, too, work is required to compress the gas; and the amount of the heat-effect produced depends only upon the amount of work done. When a substance is burnt, there are heat-effects. For instance, if carbon combines chemically with oxygen, as it does in an ordinary fire, there are heat-effects; and their amount depends only upon the loss of potential energy of the carbon and oxygen which combine.

So, in general, it may be proved that the numerical value of the heat-effect, whatever it is, depends only upon the work done; or, what is the same thing, upon the loss of energy by the bodies doing the work.

It may, therefore, be stated that heat-effects are manifestations of energy; because, in their production, energy has passed from certain bodies into those in which the "heat-effects" occur. Further, it is noticed that in every heat-effect the work done is against "forces" which act between most minute portions of matter or is concerned in some way with the motion of these minute portions. This is equivalent to saying that, in the production of heat-effects, energy becomes associated, in either a potential or kinetic form, with portions of matter which are of the general dimensions of molecules. In an exactly similar manner, it may be stated that if energy is taken away from the molecules there is what may be called a negative heat-effect, such as fall in temperature, contraction, etc. These facts will be shown more in detail in Chapter I.

Since, then, all heat-effects are caused by the passage of energy into an association with minute portions of matter, the amount of any heat-effect must be measured by the amount of energy which is involved in the phenomenon. That is, all heat-effects must be measured in ergs (Art. 68).

164. The most important sources of heat-energy are the following: —

a. Such mechanical actions as friction, compression, etc.

b. Many chemical actions, such as combustion, especially of carbon and oxygen.

c. The sun, which directly or indirectly is the chief source of the energy available for our use.

d. Electric currents, because, whenever a current passes, the temperature of the conductor is raised.

e. Changes in molecular arrangement, which will be more fully discussed in Chapter IV., e. g. freezing of water.

f. Hot-springs, volcanoes, etc., all being evidences of the high temperature of the interior of the earth.

CHAPTER I

HEAT-EFFECTS — TEMPERATURE

AMONG the most familiar heat-effects are the following:

1. Change in volume.
2. Change in temperature.
3. Change in molecular arrangement or structure, such as melting, boiling, etc.

There are also electrical and chemical effects; but these three are the most important.

165. Change in Volume. It is well known that, if any body is warmed over a fire or in other ways, its volume is changed. It in general expands. In this act of expansion, work is done in at least two ways: first, the molecules of the body are forced farther apart; second, whatever is resting on the body is moved. This first change, that of position of the molecules, evidently belongs to the third effect noted above, "Change in molecular arrangement;" and it will be discussed later. The work done by the expanding body against the matter touching it is sometimes called "external work." Thus, if a pillar which supports a building expands, it raises the building, and so does work. An ordinary solid which is surrounded by air does work in expanding by raising the atmosphere around it. So it may be said that, in expansion, energy has left the body producing the expansion; and this has been spent in two ways at least, — in overcoming some external force and in producing molecular re-arrangements. It is possible, of course, that these molecular re-arrangements may be accompanied by a *decrease* in the potential or

kinetic energy of the molecules; and in this case, instead of work being required to produce the molecular change, this change itself would furnish energy, which would be available for external work.

If the body contracts, there are, of course, molecular changes; the external forces now themselves do work; and so the body receives energy.

In any case, if there is a *uniform pressure, p,* over the entire body, and if its volume changes from V_0 to V_1, the external work done is $p\,(V_1 - V_0)$ (see Art. 95). If $V_0 < V_1$, this work is done against the external forces. If $V_0 > V_1$, the work is negative, meaning that it is done *by* the forces; and energy is given the body.

166. Change in Temperature. That heat-effect which is best known is change in temperature, because our senses allow us to tell with some degree of accuracy that when held over a fire, a body, as a rule, becomes hotter. All bodies do not, however, become hotter under these conditions. For instance, a block of ice may be heated until it all melts; and there is no change in temperature. But these questions will be discussed later. So will the more exact ideas of temperature, and the methods of measurement.

As was stated in the general introduction (Art. 5), the present theory of matter is to regard all minute portions of it as being in motion. We shall see evidences of this later on; and it will be shown also that in gases there is no question but that the temperature and the average kinetic energy of translation of the molecules are most intimately connected. If this energy is increased in any way, there is an increase in the temperature, and *vice-versa.* It is also believed that this is more or less true of solids and liquids. So that, if the work done on the body is spent in increasing the kinetic energy of translation of the molecules, the particular heat-effect manifested is rise in temperature.

167. Molecular Changes. Any alteration of molecular arrangement or structure must involve changes in energy, either potential or kinetic. Changes in kinetic energy are in general, as just noted, made evident by changes in temperature. But among so-called heat-effects there are a great many changes which are mainly changes in potential energy of the molecular structures. Fusion, or the change of state of a body from solid to liquid, is an illustration of this. In this process work must be done on the molecules, that is, energy must be added to them, so as to give them their increased freedom of motion, which they have in the liquid state. Conversely, if a liquid is solidified, energy must be taken away from the molecules; and the amount taken away must equal exactly the amount required to liquefy the solid.

Evaporation, or the change of state of a body from liquid to gas, is another illustration. So is sublimation, or the change from solid to gas directly. Any change in volume also presupposes molecular changes; and this, in general, involves changes in potential energy. So does change in shape. Dissociation is also an example, because in this phenomenon molecules are broken up into simpler parts. Combination, the reverse of dissociation, is another. So is solution of one body in another, because the molecules (or their parts) become differently arranged.

In all these cases, if there is an increase in potential energy, work is done on the molecules. This gain in potential energy may come from the fact that the molecules themselves are losing kinetic energy, and in this case the temperature falls; or, if the temperature remains constant, the energy must come from outside. If there is a decrease in potential energy, this may correspond to an increase in kinetic energy and so to a rise in temperature, or to an emission of energy to surrounding bodies.

The internal changes are, then, changes in temperature and changes in molecular arrangement or structure; and,

as a result of them, the molecular energy of the body is changed. The molecular energy of a body at any instant is called its "intrinsic energy" in that condition; and so these changes are called changes in intrinsic energy. If the body ever returns, after any cycle of changes, to its original condition, its intrinsic energy is the same as before, because it depends only upon the body itself.

When molecular energy is given a body, other effects than these may, of course, be produced, such as electrical effects; but the Conservation of Energy requires that the energy added shall exactly equal the sum of the external and the internal work done. This is perfectly borne out by all known experimental evidence. Stated in other words: energy added = change in intrinsic energy + external work done.

Temperature

168. The temperature of a body is a property of the body which expresses its thermal relation to other bodies. One body is said to have a higher temperature than another, if, when the two are placed in intimate contact, the first loses molecular energy and the second gains it. If a body is gaining molecular energy, some heat-effect is being produced; and, since heat-effects are easily observed, it may at once be seen which body has the higher temperature. If two bodies are at the same temperature, neither one gains or loses molecular energy; and this condition may be easily determined by experiment. Further, it is proved by experiment that, if two bodies have the same temperature as a third body, the two bodies themselves have the same temperature. This fact permits the comparison of the temperatures of two bodies which cannot be conveniently placed in contact, by means of a third body which can be compared with the two separately.

This third body must have some property which is easily affected by changes in molecular energy; but it is

entirely immaterial what property this is, only it must admit of exact measurement. Such a body may be called a "thermometer," because by means of it a numerical value may be given to the temperature of any body.

In the ordinary thermometer, changes in temperature are recognized by changes in volume of some fluid, e.g. mercury or alcohol or air, enclosed in glass. The alterations in volume are registered by divisions made on the glass tubes which enclose the fluid; a liquid has of course a free surface, but a gas may be separated from the outside air by a short mercury index which marks the volume. The volume of the fluid may be noted under two different definite conditions. Call these two volumes v_1 and v_2. Let it be desired to distinguish between the two given conditions n degrees of temperature. Then one degree may be defined as corresponding to a difference in volume of the fluid of $(v_2 - v_1) / n$. So, if t_1 is a number given arbitrarily to the temperature which corresponds to the volume v_1, the numerical value of the temperature which corresponds to the volume v is

$$t = t_1 + \frac{v - v_1}{\frac{v_2 - v_1}{n}} = t_1 + n\frac{v - v_1}{v_2 - v_1} \quad . \quad . \quad . \quad . \quad (1)$$

If mercury is the fluid used, the thermometer is called a "mercury-thermometer." If air is used, the instrument is called an "air-thermometer," etc. By means of such an instrument, a numerical value can be given to the temperature of any body; but the number depends upon two arbitrary quantities, t_1 and n, and upon the fluid used, and so is itself perfectly arbitrary.

By general consent the two definite conditions which fix the two standard temperatures t_1 and t_2 are the temperature of equilibrium of ice and water and of water and steam, when the corrected atmospheric pressure is 76 cm. of mercury. (See Art. 175.) Accurate experiments have

shown that these two temperatures are always the same the world over; and they are also very convenient for ordinary purposes. Further, by universal agreement among scientists, there are distinguished 100 intermediate degrees of temperature; and the temperature of equilibrium of ice and water is called 0°; i. e. $n = 100$, $t_1 = 0$. If v_0 is the volume at 0° and v_{100} that at 100°, the temperature which corresponds to any volume, v, is

$$t = 0° + 100° \frac{v - v_0}{v_{100} - v_0} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

This number is called the temperature on the "centigrade" scale with the mercury-, or air-, etc., thermometer. This scale can obviously extend below 0° and above 100°, if v is less than v_0 or greater than v_{100}.

FIG. 132.

A mercury-thermometer consists, in general, of a glass bulb with a long narrow glass tube attached, the mercury filling the bulb and part of the tube. This tube is closed at the further end; and the space above the mercury is often a vacuum, although sometimes a gas like nitrogen is put in under low pressure. The narrow tube is supposed to be perfectly uniform; and divisions are marked on it corresponding to equal changes in length. The instrument-maker generally marks 0 at the point where the mercury stands at 0°, and 100 where it stands at 100°, dividing the intervening column into 100 equal parts. Of course, any maker is liable to commit an error, and so each thermometer must be carefully standardized and calibrated. That is, it must be tested to see if, when at the temperature of equilibrium of ice and water, the reading on the scale is 0; and if, when at the temperature of equilibrium of water and steam under the pressure of the standard

atmosphere, the reading is 100. And further, it must be examined to see if the long tube is perfectly uniform, so that equal lengths correspond to equal volumes. Methods for this examination are taught in all laboratories. (It may be noted here, simply for reference, that if the atmospheric pressure is 1 cm. greater than 76 cm. of mercury, the temperature of equilibrium of water and steam rises 0.°366 centigrade; and for other small variations in pressure the changes in temperature are in proportion. If the pressure falls, the equilibrium-temperature also decreases. See Article 194.)

An ordinary thermometer is liable to many errors, all of which must be carefully taken into account. These necessary corrections are all explained in laboratory manuals, and are, in almost all cases, due to the fact that glass is a substance which, when once expanded, does not return to its previous volume after the temperature is brought back to its former value, until many days or months have elapsed.

A mercury-thermometer cannot be used, of course, except between those temperatures for which mercury remains a liquid, that is, between − 39° C. and 350° C. And, since mercury expands at different rates at different temperatures, there is no agreement between the meaning of 1° at different temperatures. A gas-thermometer is free from these limitations, because it does not liquefy except at enormously low temperatures, and because it expands nearly uniformly under constant conditions of pressure. But a gas must be enclosed in some solid; and, as this melts at some high temperature, even a gas-thermometer cannot be used at very high temperatures. A gas-thermometer is really used but rarely, except for standardizing purposes, owing to the large volume which the instrument must necessarily have.

CHAPTER II

CHANGES IN VOLUME

SOLIDS

169. Linear Expansion. Whenever a solid is held over a fire, or is otherwise heated, its volume and its temperature both, as a rule, change. (If the solid is at its melting temperature, this is not so.) If the solid is isotropic, i. e. has the same properties in all directions, the expansions of all *lines* of equal length in the body are the same, as may be proved by experiment. If the body is not isotropic, e. g. if it is a crystal, this is not true.

If a long rod of any substance has its length measured at different temperatures, there is a connection between the change in length and the change in temperature. Thus let

$$a_1 = \text{length at temperature } t_1,$$
$$a_2 = \text{“} \quad \text{“} \quad \text{“} \quad t_2;$$

then it is found by experiment that

$$a_2 - a_1 = \alpha\, a_1 (t_2 - t_1) \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where α is very nearly a constant for any one substance for different values of t_2.

If $t_2 - t_1 = 1$, i. e. if the temperature is raised one degree, $\alpha = \dfrac{a_2 - a_1}{a_1}$; and this ratio is called the "coefficient of linear expansion" of this substance at the temperature t_1. α is nearly the same at all temperatures, but not exactly.

If t_2 is not equal to $t_1 + 1$, the ratio $\alpha = \dfrac{a_2 - a_1}{t_2 - t_1} \cdot \dfrac{1}{a_1}$ is

called the "average coefficient of linear expansion between t_1 and t_2."

The figures in the following table refer to the changes in length of rods of different substances, which are first measured at 0° C.; i. e. $t_1 = 0$, and $a_1 = a_0$ is the length at 0°,

$$\alpha = \frac{a_2 - a_0}{t_2 a_0};$$

$$\therefore a_2 = a_0 (1 + \alpha t_2) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

TABLE III

AVERAGE COEFFICIENTS OF LINEAR EXPANSION BETWEEN 0° AND 100° C.

Aluminum	0.000023	Platinum . .	0.000009
Brass	0.000018	Silver	0.000019
Copper	0.000017	Steel	0.000011
Glass	0.000009	Tin	0.000023
Iron (soft)	0.000012	Zinc	0.000029
Iron (cast)	0.0000105		

Since α is different for different substances, it is possible so to combine different rods that the resulting length does not change when the temperature is altered. This is the principle of one form of compensating pendulum.

If the body is isotropic, as stated above, α is the same for all directions; but if the body is not isotropic, α has different values in different directions. In a crystal which has three different axes at right angles to one another, α has three different values corresponding to these three directions. In general α is positive, i. e. the length of a line of the body increases as the temperature rises; but this is not true in all cases: for instance, if the temperature of a stretched rubber cord is raised, it tends to contract.

The measurement of α for a solid is not difficult, as it

involves only the measurement of the length of a bar at two different temperatures. Fuller details are given in laboratory manuals. (The most accurate method depends upon the interference of light-waves, and the principle will be discussed later.)

170. Application of Principle of Stable Equilibrium. If a bar expands when its temperature is raised, stretching the bar mechanically will cool it. For a bar held in a region of constant temperature is in stable equilibrium; and in accordance with the principle of stable equilibrium (see Art. 66), if when the temperature rises expansion is produced, expansion of itself must tend to cause a fall in temperature. Compression will produce a rise in temperature.

If, on the other hand, the substance of the bar is such that raising the temperature produces contraction, stretching will produce a rise in temperature.

171. Cubical Expansion. Since each line in a solid changes when its temperature is altered, the volume will also change (unless the body is crystalline and one axis contracts or expands enough to balance the change of the other two). In an isotropic body whose coefficient of linear expansion, referred to 0° C., is a, formula (2) gives:

$$a = \frac{a_2 - a_0}{t_2 \, a_0},$$

or

$$a_2 = a_0 \, (1 + a \, t_2).$$

Consider a cube each of whose edges had the length a_0 at 0° C. Then at the temperature $t_2°$, each edge will have the length a_2. The initial volume, v_0, was a^3; the final volume, v_2, at $t_2°$ is $a_2{}^3$. Hence, cubing formula (2),

$$v_2 = v_0 \, (1 + a \, t_2)^3.$$

But a is so small for all solids that a^2 is too minute to notice. Hence

$$(1 + a \, t_2)^3 = 1 + 3 \, a \, t_2.$$

Or $$v_2 = v_0 (1 + 3\,a\,t_2).$$

That is, $$3\,a = \frac{v_2 - v_0}{t_2\,v_0}.$$

And so, corresponding to the coefficient of linear expansion a, the coefficient of cubical expansion is $3\,a$. This is sometimes written β. That is,

$$\left.\begin{aligned} \beta &= \frac{v_2 - v_0}{t_2\,v_0} \\ \text{or} \qquad v_2 &= v_0 (1 + \beta\,t_2) \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

A solid which has in it any kind of a cavity expands exactly as if there were none. For imagine the actual solid to be constructed of minute cubes (or bricks). As the temperature is raised, each cube expands into a larger cube; and the surface of any cavity bounded by these larger cubes is just as much larger than it was before as it would have been if the space had been solid and not hollow.

Liquids

172. Cubical Expansion. If the temperature of a liquid is changed, its volume is also; and the connection between these changes may be expressed as it was for a solid. If v_1 and v_2 are the volumes of a particular liquid at the temperatures t_1 and t_2, it is found by experiment that

$$v_2 - v_1 = \beta\,v_1 (t_2 - t_1), \quad . \quad . \quad . \quad . \quad . \quad (4)$$

where, for a given liquid, β is nearly a constant for various values of t_2, and is called the average coefficient of cubical expansion between the temperatures t_1 and t_2. As generally measured, β refers to the centigrade scale and to the temperature $0°$ as the starting-point. That is, $v_1 = v_0$, $t_1 = 0$.

$$\left.\begin{aligned} \text{Hence,} \qquad v_2 - v_0 &= \beta\,v_0\,t_2 \\ \text{or} \qquad v_2 &= v_0 (1 + \beta\,t_2) \\ \text{So} \qquad \beta &= \frac{v_2 - v_0}{v_0\,t_2} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Some values of β for different substances are given in table. It will be noticed that they are larger than for the ordinary solids.

TABLE IV

AVERAGE COEFFICIENTS OF CUBICAL EXPANSION OF LIQUIDS

Alcohol . .	0°–80° 0.00105	Mercury .	0°–100° 0.000182
Ethyl Ether	0°–33° 0.00210	Turpentine	9°–106° 0.00105

173. Water. β is not constant for any one liquid, but changes with the temperature, just as it does with solids. In one liquid, water, β changes greatly. In fact, between 0° and 4° C. β for water is negative; while from 4° to 100° C. it is positive. So water is densest at 4° C. as noted in Article 8. This fact is most important in the economy of nature; because, as the water in a lake or river becomes cold, the denser portions sink to the bottom; and so, while the water on the top may be 0°, that at the bottom will be only 4° C. Consequently, lakes and rivers freeze on the top, not on the bottom.

174. Application of Principle of Stable Equilibrium. A liquid kept in a region of constant temperature is in stable equilibrium. So if, when the temperature of the liquid is increased, there is expansion, then expanding the liquid will produce a fall in temperature, and compressing the liquid will cause a rise in temperature. But, if the liquid is one which contracts when its temperature is raised, then compressing it will produce a fall in temperature.

175. Barometric Correction. As explained in the section on the mercury barometer (Art. 100), the actual height of the mercury column varies with the temperature, because the density of the mercury changes; and so the observed height is "corrected" to give the height to which the column would have risen if the density of the mercury had

been that which it is at 0° C. This correction is made in this way. Let ρ and h be the density and height actually observed at $t°$ C; and ρ_0 and h_0 the density at 0° and the "corrected" height. ρ_0 and h_0 must be such that the pressure due to a column of height h_0 and density ρ_0 shall equal the existing pressure. That is,

$$\rho g h = \rho_0 g h_0,$$

or

$$h_0 = \rho \frac{h}{\rho_0}.$$

But the change in density of 1 gram of mercury when its temperature is changed from 0^0 to $t°$ C. is easily calculated. Call v_0 the volume at 0°; and v that at $t°$ C. Hence

$$1 = v_0 \rho_0 = v \rho. \quad \therefore \ \rho / \rho_0 = v_0 / v \quad . \quad . \quad . \quad (6)$$

But

$$v = v_0 (1 + \beta t),$$

where β is the average coefficient of cubical expansion of mercury between 0^0 and $t°$ C. Hence

$$\rho / \rho_0 = \frac{1}{1 + \beta t}, \quad . \quad . \quad . \quad . \quad . \quad (7)$$

and so

$$h_0 = \frac{h}{1 + \beta t} = h (1 - \beta t), \quad . \quad . \quad . \quad (8)$$

since β is so small that its square may be neglected in comparison with 1.

But h is in general measured in divisions ruled on the metal or glass case which encloses the mercury; and, if one of these divisions is 1 cm. long at 0° C. its length in cm. at $t°$ is $a = a_0 (1 + \alpha t)$, where $a_0 = 1$, and α is the average coefficient of linear expansion of the metal or glass. That is,

$$a = 1 + \alpha t.$$

Hence the true corrected height in centimetres is

$$h_0 = h (1 + \alpha t)(1 - \beta t) = h [1 - (\beta - \alpha) t] \quad . \quad . \quad (9)$$

For mercury $\beta = 0.0001816$;
For brass $\alpha = 0.000018$.

And so for an ordinary mercury barometer with a brass case

$$h_0 = h\,(1 - 0.00016\,t) \quad . \quad . \quad . \quad . \quad . \quad (10)$$

176. Measurement of β. A liquid must obviously be held in a solid; and so, when the temperature is changed, the volume of the solid changes as well as that of the liquid. Consequently, the apparent or the observed change of the liquid is the actual true change *less* the change of the solid. If a liquid is enclosed in a bulb with a fine connected tube, the volumes of equal lengths of which are known, the apparent change may be at once measured by the change in height of the top of the liquid column. So, if the volume of the bulb is known, and if it is of a substance whose coefficient of cubical expansion is known, its total expansion may be calculated; and then the true expansion of the liquid is at once given. From a knowledge of this true expansion, the change in temperature and the initial volume of the liquid, β may be calculated, because

$$\beta = \frac{v_2 - v_1}{(t_2 - t_1)\,v_1}\,.$$

FIG. 133.

This same experiment permits one to measure the cubical coefficient of expansion of the solid which encloses the liquid, if the absolute expansion of the liquid is known. This method is very commonly used, because the absolute expansion of mercury is quite accurately known.

To avoid the necessity of determining the coefficient of expansion of the enclosing solid, another method has been devised, which is independent of the nature of the solid. Its principle depends upon the fact that the height to which a liquid rises above its free surface in a tube

which dips into a basin of the liquid is entirely independent of the size or shape of the tube, and depends only upon the pressure and the density. A double U-tube is constructed with the outer open arms quite long. The liquid to be examined is poured into the outer arms, but the two portions of the liquid are not allowed to meet in the middle branch. One-half the tube is enclosed in a space whose temperature is t_1; and the other in a space whose temperature is t_2. By compressing the air in the middle portion of the tube, the liquid may be forced up to considerable heights in the two outer arms. Let the two vertical heights of the surfaces of the outer columns above the free surfaces of their respective inner columns be h_1 and h_2, and let their respective densities be ρ_1 and ρ_2. Then, since both columns have the same pressures at both their outer and their inner ends,

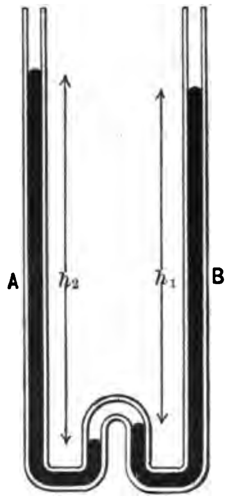


FIG. 134.

$$\rho_1 g h_1 = \rho_2 g h_2,$$

or

$$\rho_1 h_1 = \rho_2 h_2.$$

But, considering the volumes v_1 and v_2 of one gram of the liquid at the temperatures t_1 and t_2

$$1 = \rho_1 v_1 = \rho_2 v_2.$$

Hence

$$v_2 h_1 = v_1 h_2.$$

But

$$\beta = \frac{v_2 - v_1}{(t_2 - t_1) v_1}.$$

Hence

$$\beta = \frac{h_2 - h_1}{(t_2 - t_1) h_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

All the quantities in the fraction can be measured, and so β may be determined.

Gases

177. Cubical Expansion. All gases expand as the temperature rises if the pressure is kept constant, and the connection between volume and temperature may be expressed by the equation

$$v = v_0 (1 + \beta t), \quad . \quad . \quad . \quad . \quad . \quad (12)$$

where v_0 is the volume at 0° C., v is the volume at t° C., and β is the average coefficient of cubical expansion at constant pressure, referred to 0° C.

β is not alone nearly constant for any one gas; but it is also nearly the same for all gases. This is a most remarkable fact, and bears witness to the simplicity of the laws of gases. That β is the same for all gases is called Charles' Law, sometimes Gay-Lussac's. Its value is almost exactly 0.003665.

Since, in the expansion of a gas at constant pressure, the mass of the gas does not change, the above formula may be changed into

$$\rho_0 = \rho (1 + \beta t),$$

because $$v_0 = \frac{\text{mass}}{\rho_0}, \; v = \frac{\text{mass}}{\rho},$$

and so $$v_0 \rho_0 = v \rho.$$

178. Change of Pressure and Temperature. If the pressure changes, the volume does also in general. In particular, if the temperature is constant, and the pressure changes, Boyle's Law (see Art. 108) gives the connection between the pressures and volumes (or densities), viz.:

$$p v = k m,$$

or $$p / \rho = k,$$

where k is a constant for any gas at any one temperature. If both temperature and pressure change, the process may be considered as taking place in two stages, — a change in

temperature at constant pressure, then a change in pressure at constant temperature.

Let the gas have a density ρ_0 at 0° C. and pressure p_0. If the temperature is raised to t° C. at constant pressure,

$$\rho_t = \rho_0 / (1 + \beta t).$$

Now, if the pressure is changed from p_0 to p, the temperature being constant, the density will be changed from ρ_t to ρ, where

$$p_0 / \rho_t = p / \rho.$$

Hence
$$p / \rho = p_0 (1 + \beta t) / \rho_0,$$

or
$$\frac{p}{\rho (1 + \beta t)} = \frac{p_0}{\rho_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

Consequently, the resulting density, ρ, at the temperature t° C., and the pressure, p, may be calculated. Or, conversely, if the density is known at any temperature and pressure, the density at 0° C. and at any pressure, p_0, may be deduced. In chemistry the volume of a gas "under standard conditions" is the volume which a given mass of the gas would occupy at 0° C., and at a pressure equal to 76 cm. of mercury. Thus, substituting $v / v_0 = \rho_0 / \rho$,

$$v_0 = \frac{p v}{p_0 (1 + \beta t)}, \quad . \quad . \quad . \quad . \quad . \quad (14)$$

where v is the observed volume at pressure p and temperature t° C.; and v_0 is the "volume under standard conditions," if p_0 is pressure of 76 cm. of mercury. p / p_0 is the ratio of the actual pressure of the gas measured in centimetres of mercury to 76.

Another deduction from the general formula is that, if the volume and mass of a gas are kept unchanged, but the temperature varied, i. e. if $\rho = \rho_0$, the pressure changes so that

$$p = p_0 (1 + \beta t) \quad . \quad . \quad . \quad . \quad . \quad (15)$$

That is, the pressure at constant volume changes with the temperature at the same rate that the volume does at constant pressure.

The value of β as found by experiment for gases is nearly 0.003665; so that $\frac{1}{\beta} = 273$ nearly. Substituting this value in the formula

$$\frac{p}{\rho\,(1+\beta t)} = \frac{p_0}{\rho_0},$$

it becomes

$$\frac{p}{\rho\,(273+t)} = \frac{p_0}{\rho_0\,273}.$$

But for any one gas p_0 / ρ_0 is a constant; for ρ_0 is the density of the gas at 0° C. when the pressure is p_0; and by Boyle's Law p_0 / ρ_0 is constant for a given temperature. Call $\frac{p_0}{\rho_0\,273} \equiv R$. It is evidently a constant which can be easily found by experiment for any gas, and is sometimes called the "gas-constant" for that gas. Then

$$\frac{p}{\rho\,(273+t)} = R.$$

$273 + t$ is a number which gives the reading which the temperature $t°$ on the centigrade scale would have, if the scale started from 273° C. below 0° C. This number; $273 + t$, is sometimes written T, and is called the "absolute" gas-thermometer temperature. It is evidently a constant for all gases.

The formula then becomes

$$\frac{p}{\rho\,T} = R \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

or, substituting $\rho = m / v$,

$$\frac{p\,v}{T} = R\,m \quad . \quad . \quad . \quad . \quad . \quad (16\,a)$$

This is the fundamental law for a gas; and from it all laws may be derived. It is based upon two experimental laws, Boyle's and Charles'.

[It is evident that if $T \doteq 0$, either p or v must equal 0; but this has no physical meaning. In fact, the law applies only to gases; and there is no reason to believe that at $-273°$ C. matter would exist in the form of a gas. The law is not *exactly* true of any actual gas, but expresses the ideal behavior of a "perfect" gas.]

179. Measurement of β. The constant β may be measured in two ways: the pressure may be kept constant and the volume measured at different temperatures, in which case

$$v = v_0 (1 + \beta t),$$

or, the volume may be kept constant and the change in pressure measured when the temperature changes, in which case

$$p = p_0 (1 + \beta t).$$

With all actual gases the values of β as found by these two methods differ slightly, showing that Boyle's and Charles' laws are not rigidly true of actual gases.

Application of Principle of Stable Equilibrium. A mass of gas enclosed in a vessel (e. g. a cylinder with movable piston) is in stable equilibrium if it is in a region of constant temperature. All gases expand if the temperature is raised; so expanding a gas lowers the temperature, compressing it raises the temperature.

CHAPTER III

CHANGE OF TEMPERATURE

180. WHEN molecular energy is taken away from a body or added to it, there is in general what is called a change in temperature; but for equal amounts of energy the temperature is affected differently for different bodies.

Calorie. To investigate this difference, it would be necessary to measure the change in temperature of different bodies when equal amounts of energy are given their minutest portions so as to produce heat-effects, or to determine the amount of energy necessary to raise the temperature one degree. This would be a most difficult process, because our ordinary methods of producing changes in temperature do not admit easily of measurement in terms of ergs, the unit of energy. The best method is to take some secondary unit, in terms of which the energy going into heat-effects may be easily measured, and whose value in terms of ergs may be once for all determined. Such a unit is the energy required to raise the temperature of one gram of water from 10° C. to 11° C.; and it is called a "calorie" or a "thermal unit at 10° C." Its value in terms of ergs is found by experiment to be 4.2×10^7; and it is evidently a most convenient practical unit, in terms of which to measure the energy required to produce in any body rise in temperature or any heat-effect. The amount of energy required to raise the temperature of one gram of water one degree anywhere else than at 10° C. is not 1 calorie; but it differs from that so slightly that in ordinary experiments the difference may be neglected.

181. Specific Heat. The number of calories which is required to raise the temperature of 1 gram of any substance from $t°$ C. to $(t + 1)°$ C. is called the "specific heat" of that substance at the temperature $t°$. The specific heat of any substance is different for different temperatures; and so the average specific heat is the number of calories required to raise the temperature of 1 gram a certain number of degrees centigrade, divided by the number of degrees. The specific heat varies greatly for different substances. Some few values are given in the following table:—

TABLE V

AVERAGE SPECIFIC HEATS

Alcohol	....	0.615	Mercury	20°–50°	0.0333
Aluminum	0°–100°	0.2185	Paraffin	.....	0.683
Brass	....	0.09	Platinum	0°–100°	0.0323
Copper	0°–100°	0.0933	Silver	0°–100°	0.0568
Glass	....	0.20	Tin	0°–100°	0.0559
Iron	0°–100°	0.1130	Turpentine	.....	0.467
Lead	0°–100°	0.0315			

It is evident, of course, that if c is the specific heat of any substance, the number of calories required to raise the temperature of m grams 1 degree C. is $m\,c$. And, if c is the average specific heat between the temperatures $t°$ and $T°$, the number of calories required to raise the temperature of m grams that amount is $m\,c\,(T° - t°)$. Conversely, if m grams cool from $T°$ to $t°$, that number of calories must leave the body, if it returns to its original condition.

Some of the energy which is added to the body when the temperature is raised, is spent in producing molecular changes which involve changes in potential energy (except in perfect gases) and also in doing external work, so the *true* specific heat would be only that portion of the energy

which is spent in raising temperature. It is, however, practically impossible to measure this quantity except in the case of gases.

182. Specific Heat of Gases. It will be shown later (see Art. 186) that in the more perfect gases, such as hydrogen and oxygen, practically no energy is required to produce any rearrangement of their molecules. If the volume of a gas is then increased, the only energy required is that necessary to do the external work, provided the temperature does not change. The number of calories required to raise the temperature of 1 gram of a gas 1° C., if the volume is kept constant, is called the "specific heat at constant volume." It is ordinarily written C_v; and it is the true specific heat of the gas. The number of calories required to raise the temperature of 1 gram of the gas 1° C. when the volume increases, but the pressure is kept constant, is called the "specific heat at constant pressure," and is ordinarily written C_p. C_p is, of course, greater than C_v; and the difference $C_p - C_v$, expressed in ergs, is, by what was said above, the amount of external work done when 1 gram of the gas expands at constant pressure owing to its temperature being raised 1° C. This may be easily calculated. By Formula 16 *a*, Chapter II.,

$$p\,v = R\,m\,T.$$

So that, if we deduce the change in volume of 1 gram when the temperature is raised from T to $T + 1$, the pressure being constant, we have

$$\begin{aligned} p\,v &= R\,T \\ p\,v_1 &= R(T+1) \\ \therefore\ p\,(v_1 - v) &= R. \end{aligned}$$

But the external work done by the gas is (by Art. 104) the product of the pressure and the change in volume, $p\,(v_1 - v)$, i. e. R.

Call $J \equiv 4.2 \times 10^7$. Hence the number of ergs in $C_p - C_v$ calories is $J(C_p - C_v)$. Consequently,

$$J(C_p - C_v) = R \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

All four of these quantities may be measured, and the formula is found to be verified.

It may be well to add that for any gas both C_p and C_v are nearly constant for all temperatures; and so their ratio is also. But both C and C_v change with the pressure; that is, if the gas is compressed, these quantities change. For air and carbon dioxide (CO_2) C_v increases if the pressure is increased; while for hydrogen just the opposite is the case. The quantity C_p / C_v is commonly called γ; and it may be found by direct experiment. (See Art. 119.) Further, as noted in Article 109, the ratio of the elasticity of a gas when it is compressed very rapidly to the elasticity when the gas is compressed slowly may be proved by both theory and experiment to be the same as the ratio C_p / C_v.

Some values of C_p, C_v, and γ are given in the following table; and it will be seen that γ as found by experiment very nearly equals C_p / C_v.

TABLE VI

Specific Heats of Gases

	C_p	C_v	γ
Air	0.237	0.171	1.40+
"Argon"	. . .	. . .	1.66
Chlorine	0.124	. . .	1.33
CO_2	. . .	.165	1.29
"Helium"	. . .	. . .	1.74 (?)
Hydrogen	0.340	.241	1.41
Mercury (vapor) . . .	. . .	. . .	1.67
Nitrogen	0.244	. . .	1.41
Oxygen	0.217	. . .	1.41−

183. Dulong and Petit's Law. It has been noticed by several observers that there is a simple mathematical connection between the specific heats of various substances and their atomic weights. (See some text-book on Chemistry.) One of these relations was first observed by Dulong and Petit. They found that if the product of the specific heat of any elementary substance and its atomic weight is formed, the number is nearly the same for all substances. Thus

	Specific Heat.	Atomic Weight.	Product.
Copper	0.0933	63.5	6.1
Iron	0.113	54.5	6.2
Lead	0.0315	207.0	6.4
Silver	0.0568	108.0	6.2

This law is equivalent to stating that the number of calories (or the amount of energy) required to raise the temperature of one atom of any substance 1° C. is the same for all substances. An exact agreement in all the products was not to be expected, because, as stated before, the specific heat of any substance is different for different temperatures; and so it is impossible to choose those numerical values which are exactly comparable.

Other laws have been noticed for various series of compounds, and also for various gases.

184. Measurement of Specific Heat. There are several general methods for the measurement of specific heats. The best are the following: —

a. Method of Mixtures. The principle of the method is to put into a vessel of water the substance whose specific heat is desired, and to note the changes in temperature. The initial temperatures of the water and the substance are different; but their final temperature is, of course, the same.

Let M = mass of water,
m' = mass of vessel which contains the water,
c' = specific heat of this vessel,
m = mass of substance whose specific heat is desired,
c = the unknown specific heat of this substance,
T = initial temperature of the water,
t = initial temperature of the substance inserted,
T_0 = final temperature.

A number of calories equal to $m\,c\,(t - T_0)$ is taken away from the substance inserted; a number $M\,(T_0 - T)$ is given the water, and $m'\,c'\,(T_0 - T)$ to the vessel which contains the water. If no energy has escaped, the number of calories lost by the one body must equal that gained by the other two. That is,

$$m\,c\,(t - T_0) = (M + m'\,c')\,(T_0 - T) \quad . \quad . \quad . \quad (2)$$

The temperatures and the masses may be easily measured; so may $m'\,c'$, accurately enough for this experiment. Consequently c may be determined.

$m'\,c'$ is sometimes called the "water-equivalent" of the calorimeter, or the vessel containing the water. It may be determined roughly, by pouring into the empty vessel a known mass of water at a known temperature, and noticing the change in temperature.

Let M = mass of water poured in,
T = temperature of the water as it is poured in,
t = initial temperature of the vessel,
T_0 = final temperature.

Then $$M\,(T - T_0) = m'\,c'\,(T_0 - t),$$

or $$m'\,c' = M\frac{T - T_0}{T_0 - t} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

and so it may be measured.

This method for the determination of the specific heat is applicable to any solid or liquid which is not acted upon

by the water, and which may be secured in considerable amounts. If water does act chemically upon the solid or liquid, some other liquid of known specific heat may be used.

In order to avoid losses of energy by radiation or conduction, the calorimeter is polished and carefully separated from surrounding bodies by some non-conducting substance, such as cork, air, or feathers; and, further, the attempt is made so to choose the relative amounts of water and the substance inserted that $(T_0 + T) / 2$ shall equal the temperature of the surrounding air; that is, the initial temperature of the water is just as far below the temperature of the air as the final temperature is above it.

b. Method of Fusion of Ice. Experiments to be discussed later (Art. 189) prove that, in order to melt 1 gram of ice at 0° C. into water at 0°, 80 calories are required. So that, if the substance whose specific heat is desired is placed while hot in a cavity of ice or is surrounded by ice, and if the amount of ice melted can be measured, the specific heat may be determined.

Let m = mass of substance whose specific heat is desired,
M = mass of ice melted,
$t°$ = initial temperature of substance.

The substance loses $m\,c\,t$ calories, since it is cooled to 0° C.; 80 M calories are added to the ice. So, if there is no loss by radiation or otherwise,

$$80\,M = m\,c\,t$$

and

$$c = \frac{80\,M}{m\,t} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Suitable forms of apparatus for the use of this method have been devised; and it admits of great accuracy for even small amounts of liquids and solids. The best form of apparatus is Bunsen's ice calorimeter.

c. Method of Condensation of Steam. It is known, as the result of careful experiments, that 536 calories have to be taken away from 1 gram of steam at 100° C., in order to make it condense into water at 100°; and the number required at other temperatures is also known. So, if a substance at a temperature below 100° C is placed in a closed chest into which steam at 100° is suddenly admitted, there will be a condensation of steam into water on the substance until its temperature is raised to 100° C.; and, if the condensed water can be caught in a pan and weighed, the number of calories taken away from the steam can be calculated. The substance is held in some pan or vessel whose mass and specific heat must be taken into account.

Let m = mass of substance whose specific heat is desired,
c = its specific heat,
m' = mass of pan or vessel,
c' = its specific heat,
M = mass of steam condensed,
$t°$ = initial temperature of substance and pan.

Then, if the temperature is raised to 100° C.,

$$536\,M = (m\,c + m'\,c')\,(100 - t) \quad . \quad . \quad . \quad . \quad (5)$$

$m'\,c'$ may be found by a preliminary experiment in which there is no substance on the pan; so c itself may be determined.

This method is perhaps the most accurate now in use; and it can obviously be applied to any solid, liquid, or gas, because they can be all enclosed in some hollow closed vessel, like a sphere, which rests on a pan. With a gas, this method gives, of course, C_v, the specific heat at constant volume, because the slight expansion of the containing vessel can be regarded simply as a correction to be determined by a preliminary experiment.

d. Method of Flow. In this method, which is used to measure C_p, the specific heat of a gas at constant pres-

sure, a gas is forced very slowly through a long tube which is bent into two spirals; one spiral is kept in a vessel, *I*, at a constant temperature; the other is in a second vessel, *II*, containing water (or some other liquid). The gas in passing through the first spiral assumes the

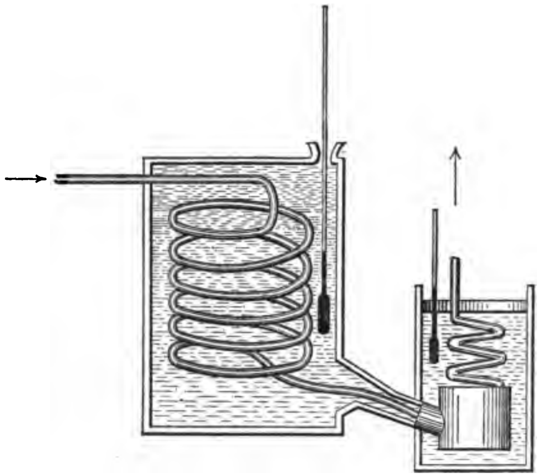


FIG. 135.

temperature of vessel *I*; it then enters the second spiral, and slowly raises the temperature of the surrounding water in vessel *II*. When it escapes from the spiral, it will have the temperature of the surrounding water, if it is drawn through slowly enough. Thus, by noting the temperature in vessel *I*, how much gas has passed through, how much water is in vessel *II*, and its changes in temperature, it is possible to calculate C_p for the gas.

CHAPTER IV

MOLECULAR CHANGES

As stated in the Introduction and Chapter I., among the most common heat-effects are those where the potential energy of the molecules is changed. Several illustrations will be discussed in this chapter.

Expansion

185. It requires the addition of energy in general to make the molecules of any body change their relative positions; and, if after the change the molecules return to their former positions and condition, the energy which they give out equals that which they received when the first change was produced. So, when any body expands, some energy is, as a rule, required to do the internal work.

186. Internal Work in a Gas. This is always true in the case of solids and liquids, but is not so for gases; or, at least, the energy required is very minute. This fact is sometimes called Joule's Law; because Joule, although not the first to prove the statement, yet was the first to show its great importance and to test it most carefully. Joule placed in a bath of water two hollow metal vessels, connected by a tube with a closed stop-cock; one vessel contained a gas at a high pressure, the other was comparatively empty. He observed that, when the cock was opened, thus allowing the gas to expand from one vessel into the other, there was no change in the average temperature of the water-bath. This proved that no heat-

energy left the gas or was added to it. In other words, no energy was required to make the gas expand. It will be noticed that the gas did no external work; the only change was one of volume. Of course, if, as a gas expands, it does external work, its temperature will fall. And in Joule's experiment the temperature of the gas which remains in the vessel where the gas was under pressure is found to be lower after the expansion, because it has done work in driving the rest of the gas out into the other vessel, where the temperature is found to be increased. But the loss of energy of the one portion exactly equals the gain of the other, and the average temperature of the entire gas is not changed, proving, as just stated, that changes in the intrinsic energy of a gas are manifest as changes in temperature.

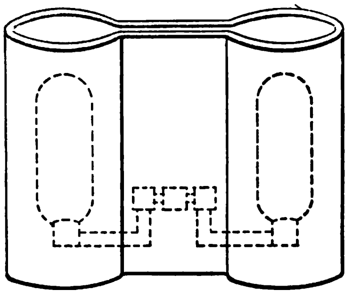


FIG. 136.

More accurate experiments show that in all gases minute changes in temperature are produced when there is expansion; but these need not be taken into account in any ordinary calculation. Hydrogen is slightly heated on expansion; all other gases are cooled.

The above statement as to the nature of a gas is equivalent to saying that its intrinsic energy (see Art. 167) depends on the temperature alone, not on its volume or pressure.

FUSION

187. Fusion is the name given to the process by which a solid passes into the liquid condition; the reverse process is called solidification. These processes obviously involve molecular changes; and, just as energy must be

added to produce fusion, so the molecules lose energy in solidification.

Fusion-Point. Most substances, under standard conditions, will begin to fuse or to solidify at a definite temperature. This is true of all crystals. If heat-energy is applied to such a solid, its temperature will continue to rise until a certain definite temperature is reached, and then further addition of heat-energy produces no change in temperature, but the solid melts. This definite temperature is known as the "fusion-point" of that substance; and, as long as the process of fusion continues, the temperature does not change. If the liquid is cooled until it begins to solidify in crystals, the temperature at which this process begins is the same as that at which the crystal fused, viz. the fusion-point; and, until the process ends and all the liquid is solidified, the temperature does not change. (Of course, this temperature is that of the entire mixture of liquid and solid, which must be well stirred.) At this temperature, then, of the fusion-point, the solid and its liquid are in equilibrium together. As long as the pressure does not change, the fusion-point remains the same. So it may be said that the solid and liquid are in equilibrium together at a definite temperature which depends only upon the pressure.

To start the solidification of a liquid, a nucleus of some kind is often necessary, or a crystal of the same form as that of the solid which is to be formed. Shaking will sometimes hasten the process by helping the molecules to rearrange themselves.

Some solids, especially the waxes, do not remain at the same temperature during the process of fusion, and the temperature at which fusion begins is not that at which solidification does.

Determination of Fusion-Point. Various methods for the determination of the fusion-point of any solid are given in laboratory manuals; but the principle made use of in all

is to determine that temperature at which the solid and the liquid are in equilibrium together. In the case of those solids like paraffin, whose fusion- and solidification-temperatures are different, the average of these two temperatures is often taken as the fusion-point or "melting-point."

Application of Principle of Stable Equilibrium. If a solid and its liquid are in contact at their fusion-point, e. g. ice and water at 0° C., they are in stable equilibrium. In all known cases solidification results in the production of heat-energy; therefore, if heat-energy is directly applied, there will be fusion, the opposite of solidification.

188. Change in Volume. When any solid fuses, there is always a change in volume. Some substances expand on fusion, e. g. wax, bismuth; while others contract, e. g. ice, cast-iron, brass. If a solid floats on its liquid, it contracts on fusion, because the density in the solid state is less than in the liquid state. While, if a solid sinks in its liquid, the opposite is true. It is obvious that sharp castings can be taken with only those substances which expand on solidification.

Effect of Pressure on Fusion. It was stated in Article 187 that under standard conditions the fusion-point was a definite temperature for any one solid. If the pressure changes, though, the fusion-point is not the same. The pressure which is required to produce even the least measurable change in the fusion-point is, however, enormously great for ordinary solids; and so the effects of ordinary atmospheric changes in pressure are too minute ever to be noticed. An increase in pressure of one atmosphere produces a change of 0.°0075 C. in the fusion-point of ice.

At great pressures, however, changes are produced; and there is, naturally, a connection between the change in volume on fusion and the effect of the increased pressure. If a solid expands on fusion, an increase in pressure would tend to prevent the change; and so the temperature at

which fusion begins is higher than it is normally, i. e. the fusion-point is raised. If, on the other hand, the solid is one which contracts on fusion, the increase in pressure helps the process, and the temperature to which it is necessary to raise the solid before it melts is lower than it is normally, i. e. the fusion-point is lowered. Therefore, since ice contracts on fusion, its melting-point is lowered by an increase in pressure. So, if two pieces of ice, whose surfaces are at 0° under ordinary conditions, are forced together with a great pressure (e. g. at two points), the melting-point is lowered, the ice melts at the point where the pressure is, and the water thus formed yields to the pressure, and flows away. But its temperature is lower than 0°C., because to produce fusion heat-energy is taken from neighboring points and added to the ice, and so it almost instantly freezes again. This process is known as "regelation;" and it explains at once the motion of glaciers over steep and rocky tracks, the formation of snow-balls under the pressure of the hands, etc.

189. "Latent-Heat of Fusion." As stated above, energy must be added to a solid in order to make it fuse, and the number of calories required to make 1 gram of any substance fuse at a fixed temperature is called the "latent-heat of fusion" of that substance at that temperature. This energy does work in producing molecular rearrangements; and, if a liquid solidifies, it emits an equal amount of energy. (A tub of water is sometimes placed in a conservatory on a night when a frost is expected, because, as the water freezes, it gives up heat-energy to the surrounding room.)

The latent-heat of fusion of any substance is generally determined by using the method of mixtures, described for measuring specific-heats, although the method of condensation of steam may be used.

The following table gives the fusion-points and the latent-heats of fusion of a few substances: —

TABLE VII

	Fusion-Point.	Latent-Heat of Fusion.
Copper	1050° C.	. . .
Ice	0°	80.
Iron	1400°–1600°	23–33
Lead	325°	5.86
Mercury	–39°	2.82
Sulphur	115°	9.37
Zinc	415°	28.

190. Effect of Dissolved Substances. When there is any substance dissolved in the liquid, the temperature of solidification is changed; a lower temperature is required than would be for the pure liquid. In general, when the liquid solidifies, it does so leaving behind it the impurity. But in certain cases, e. g. common salt, NaCl, in water, for a definite percentage of salt in solution, there is a definite temperature at which this solution solidifies as a solid of salt and ice together. Such solutions are called "cryohydrates."

When common salt is put on ice or snow, there is fusion generally, because the surrounding temperature is not low enough to keep the mixture of ice and salt in a solid form; so the ice melts and the salt is dissolved until the temperature is lowered to that of equilibrium of salt and ice in the liquid and solid conditions. In each of these processes, the melting of the ice and the solution of the salt, energy is taken away from surrounding bodies; and so their temperatures are lowered. Mixtures of ice and common salt are therefore called "freezing mixtures."

There is an intimate connection between the nature and quantity of the substance dissolved in the liquid, the liquid itself, and the lowering of the fusion-point; but the laws are too complicated to be given here.

EVAPORATION

191. Evaporation is the name given to the process by which a liquid passes into the gaseous condition. When liquids evaporate, they are generally said to form "vapors," although it is quite impossible to give any definition which will distinguish a vapor from a gas.

Vapors. Unlike the process of fusion, which does not begin until the temperature of the solid is raised to a definite degree, the process of evaporation is going on all the time from the surface of a liquid, no matter what its temperature is. If a vessel of any liquid is placed in a closed space (e. g. under a bell-jar), it is observed that the liquid evaporates gradually, but that finally the process stops; that is, the mass of the liquid remains constant. When this limit is reached, the liquid and its vapor are in equilibrium; and the vapor is said to be "saturated." Before this final stage is attained, the vapor is called "unsaturated."

192. Laws of Saturated Vapor. Unsaturated vapors are simply imperfect gases; and they obey approximately the laws of Boyle and Charles, $\frac{pv}{T} = Rm$. But saturated vapors are quite different, because the moment such a vapor is compressed some vapor condenses; and, if it is expanded, it ceases to be saturated, unless it is over a vessel of the liquid, in which case more liquid will evaporate so as to make the vapor saturated. This saturated vapor exerts a certain pressure; and it is proved by experiment that the process of evaporation comes to a state of equilibrium at a definite pressure for any given temperature. In other words, if the pressure is constant, there is a definite temperature at which a liquid and its vapor will be in equilibrium; and, conversely, for a given temperature there is a definite pressure of equilibrium. Stated in a different way, the pressure of saturated vapor depends only upon the temperature; and conversely.

Equilibrium over Curved Surfaces. It should be stated that for a given temperature the pressure of saturated vapor depends also upon the shape of the surface of the liquid with which it is in contact. There is a definite pressure for a plane surface; the pressure for a concave surface is less than this; that for a convex surface, greater.

If a liquid is placed in a closed space, and if a capillary tube is placed in the liquid, the liquid will either rise or fall, and there will finally be equilibrium between the vapor and the surfaces of the liquid. If the liquid rises, i. e. if the surface in the tube is concave, the pressure of the vapor on it is *less* than that on the plane surface of the rest of the liquid by an amount $\rho g h$, where ρ is the density of the vapor. If the liquid falls, i. e. if the surface is convex, the pressure of the vapor is *greater* than that on the plane surface by this amount. But h, the rise or fall, varies inversely as the radius of the curved surface (see Art. 97). And so these variations in vapor-pressure are very slight unless the radius of curvature is extremely minute.

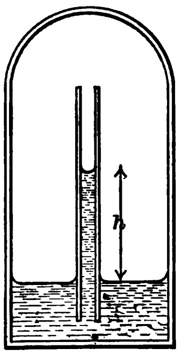


FIG. 137.

This law may be stated in another way: the pressure which is required to hold a drop of a liquid in equilibrium, to keep it from evaporating, is greater than that required to keep a plane surface in equilibrium. Further, the less the radius, the greater is the pressure required. So that in a cloud of drops of water of different sizes, the smaller drops will evaporate and disappear unless the pressure is very great; and the larger drops will grow in size.

There are two experimental methods of studying the laws of equilibrium of saturated vapor and its liquid,—the "statical" and the "dynamical."

193. The Statical Method. Some liquid is carefully forced into a mercury barometer which dips into a deep bath of mercury. The liquid rises to the top of the mercury column; and some evaporates, producing a certain pressure, and so forcing the mercury down. The distance the mercury falls marks the pressure of the vapor. Or, if H is the height of the barometer before the vapor is inserted, and h the height of the column of mercury over which the vapor is, the pressure of the vapor is $\rho g\ (H-h)$, where ρ is the density of mercury. (Corrections must be applied for the effects produced by the weight of the *liquid* and the variations in the temperature of the mercury.) The temperature of the vapor may be kept constant by surrounding the upper end of the closed tube with a bath; and the temperature of this bath may be changed from time to time so as to study the effect of different temperatures.

FIG. 138.

Two facts can now be observed. If the temperature is kept constant, the pressure of the vapor remains the same, however the tube is raised or lowered in the mercury bath so as to vary the volume. Thus, if the volume is increased, more liquid evaporates; while, if the volume is decreased, some vapor condenses. Also, if the temperature is changed, the vapor-pressure varies. In the case of all liquids at ordinary temperatures, an increase in temperature requires an increased pressure in order to keep the vapor and liquid in equilibrium, and a decrease in temperature requires a decreased pressure. Thus, if the temperature is raised, more liquid evaporates; if the temperature falls, some vapor condenses.

So, evaporation can be produced either by raising the temperature of the liquid and vapor, or by increasing the volume open to the vapor. Similarly, condensation can be

produced by lowering the temperature or by decreasing the space in which the vapor is. Illustrations of this principle are furnished in the deposition of dew, the condensation of steam in the "condenser" of a steam-engine, the condensation of vapors in general in various forms of condensers such as Liebig's. Condensation, of course, is assisted by the presence of points or nuclei around which the drops of liquid may form, as was explained in the section on Capillarity (see Art. 97).

194. The Dynamical Method. If the temperature of a liquid is raised enough, bubbles of vapor are formed throughout the liquid, especially at the surface of the vessel where it is being heated. This process is called "ebullition," or "boiling." It is assisted by the presence in the liquid of nuclei or particles of dissolved gases, or by any roughnesses on the surface of the vessel, because this helps the formation of the capillary surfaces of the bubbles. The vapor in these bubbles is obviously saturated, and the pressure of this vapor must be equal to (or *slightly* greater than) the pressure on the surface of the liquid, if the boiling is going on freely. A thermometer may be hung above the liquid and immersed in the vapor so as to measure its temperature, as shown in the figure. The pressure on the surface of the liquid may be measured, and it may be varied at will by exhausting or compressing the air above the liquid. It is

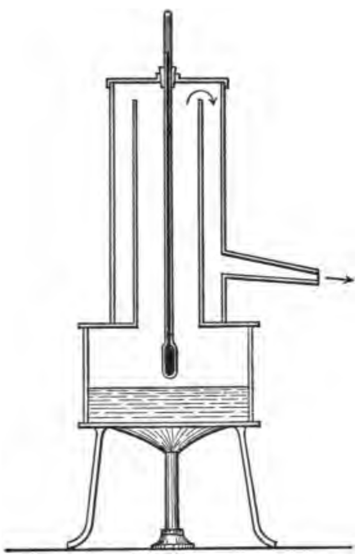


FIG. 139.

found on observation that for a definite pressure there is a definite temperature at which boiling begins, and which the vapor maintains as long as the process continues. This temperature is called the "boiling-point." In other words, the temperature of saturated vapor, i. e. of equilibrium between the liquid and its vapor, is a constant for a given pressure. If the pressure is increased, so is the boiling-point, and *vice versa*.

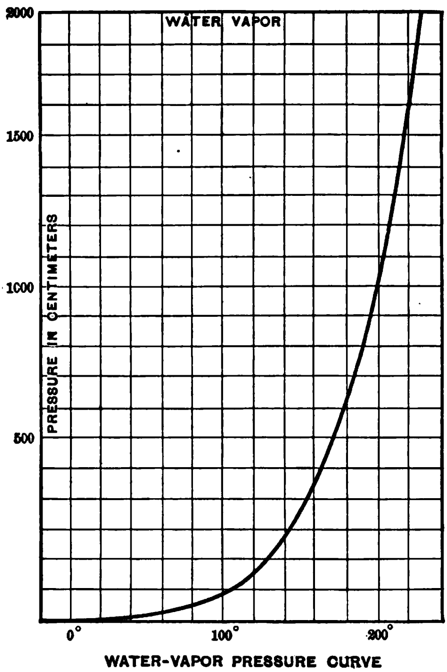


FIG. 140.

The pressure which corresponds to a given temperature is found to be the same, whether determined by the statical method or the dynamical. The connection between corresponding temperature and pressure may be shown best by a curve, called a "vapor-pressure curve," where distances

in a horizontal direction give the temperature in centigrade degrees, and vertical distances give the pressure in centimetres of mercury. The curve as shown is for water-vapor or steam.

Water boils at 100° C. when the pressure is 76 cm. of mercury; and, for slight variations of pressure, it is found, as noted in Article 168, that a difference of pressure of 1 cm. produces a difference in the boiling-point of 0°.366 C.

The following table gives the vapor-pressures of steam at different temperatures: —

TABLE VIII

VAPOR-PRESSURE OF WATER-VAPOR

Temperature, Centigrade.	Pressure in centimetres of Mercury.	Temperature, Centigrade.	Pressure in centimetres of Mercury.
−5°	0.316	98°	70.713
0°	0.457	99°	73.316
5°	0.651	99°.2	73.846
10°	0.914	99°.4	74.380
20°	1.736	99°.6	74.917
30°	3.151	99°.8	75.457
40°	5.486	100°.	76.000
50°	9.198	100°.2	76.547
60°	14.888	100°.4	77.100
70°	23.331	100°.6	77.650
80°	35.487	100°.8	78.207
90°	52.547	101°	78.767
95°	63.600	102°	81.609
97°	68.188	110°	107.54

195. Latent Heat of Evaporation. Energy is of course required to produce evaporation; and, conversely, if condensation takes place, energy is given out. The fact that energy is taken from surrounding objects when a liquid

evaporates is familiar to every one by the cold sensation felt when water evaporates from one's hand or body. Water may be frozen as a consequence of rapid evaporation; and extremely low temperatures may be secured by producing evaporation from various liquids, which themselves boil at low temperatures.

The number of calories required to evaporate 1 gram of any liquid at a given temperature is called the "latent heat of evaporation" of that liquid at that temperature. It is usually determined by a modification of the method of mixtures (see Art. 184). A known number of grams of the vapor are conducted at a known temperature into a known number of grams of the liquid, and the rise in temperature of the liquid is noted.

The following table contains the latent heats of evaporation of some liquids and their boiling-points, at standard atmospheric pressure: —

TABLE IX

	Boiling-Point.	Latent Heat.
Alcohol (ethyl)	78°.4 C.	209
Carbon Dioxide	–80°	72 at –25°
Chloroform	61°.20	58.5
Cyanogen	–20°.7	103 at 0°
Ether (ethyl)	34°.9	90
Hydrogen	–243°	
Mercury	357°	62
Oxygen	–184°	
Water	100°	535.9

Application of Principle of Stable Equilibrium. A liquid and its saturated vapor are in stable equilibrium at a definite temperature and pressure. Condensation of vapor always develops heat-energy; consequently, if heat-energy

is added, the opposite of condensation must be produced, that is, evaporation.

196. Dalton's Law for Mixtures of Vapors. If two vapors or gases are enclosed in the same space, it has been found by experiment that their mixture is uniform, and that each vapor or gas produces (nearly) the same pressure as it would by itself if the other were not there, provided that the two gases or vapors do not act on each other chemically. This law is sometimes called "Dalton's Law of Mixtures."

197. Atmospheric Moisture. There is always some water-vapor in the atmosphere; and, if the temperature is lowered enough, the vapor becomes saturated, and then for a slightly lower temperature dew is formed. This temperature is called the "Dew-Point." The "hygrometric state" at any temperature of the air is the ratio of the actual amount of moisture in the air to the amount that the air could possibly hold at the existing temperature. Its value may be easily found by experiment, but the method is not important enough for discussion here.

198. Spheroidal State. If a small amount of liquid is dropped on a very hot metal plate, it may be observed that the liquid does not touch the plate; it is supported on a cushion of its own vapor which is produced by the hot plate. This condition is called the "spheroidal state" from the shape which the drop of liquid assumes. The plate must of course be extremely hot, e. g. red-hot silver or platinum.

For a similar reason a hand moistened with water may be dipped with impunity into a pot of melted lead, if the temperature of the lead is very hot.

199. Effect of Dissolved Substances. If any substance is dissolved in a liquid, the boiling-point of the solution is raised; that is, the temperature of equilibrium corresponding to any pressure is increased. Consequently the pressure which corresponds to any temperature is lowered. The vapor formed over the solution is, as a rule, the vapor of

the pure liquid. Many interesting facts and laws have been observed as to the nature of solutions and their boiling-points and vapor-pressures; but they cannot be discussed here.

200. Liquefaction of Gases. All known gases with one or two exceptions, have been liquefied, and there is no reason why these gases may not be liquefied by suitable means. The method used is to lower the temperature of the gas as far as possible and then to compress it. It has been shown that, unless the temperature is low enough, no amount of compression will ever liquefy the gas. There is, in fact, for each gas a definite temperature, known as the "critical temperature," below which the gas must be before it can be liquefied. Some critical temperatures are given in the following table: —

TABLE X

CRITICAL TEMPERATURES

Alcohol . .	243°.6 C.	Hydrogen .	–234°
Ammonia . .	130°	Nitrogen . .	–146°
"Argon"	–121°	Oxygen . .	–119°
CO_2 . . .	31°	SO_2 . . .	156°
Chloroform .	260°	Water . . .	365°.

Isothermals. The entire process of liquefaction may best be studied by a consideration of certain curves which can be drawn to represent the behavior of liquids and vapors.

An "isothermal" is a curve which is the locus of points representing the condition of a given body while its temperature is kept constant, but its other properties varied. Usually the varying properties considered are the pressure and volume; and the curves are drawn so that distances in a horizontal direction represent volumes; and in a vertical one pressures. The isothermal for a perfect gas is as shown;

because, for a gas at constant temperature, $pv = \text{const.}$ And the isothermal for a saturated vapor is a horizontal line; because at constant temperature the pressure is constant, being independent of the volume. For a liquid, the isothermal is a line nearly vertical, because, in order to produce even a slight decrease in volume, an enormous increase in pressure is necessary.

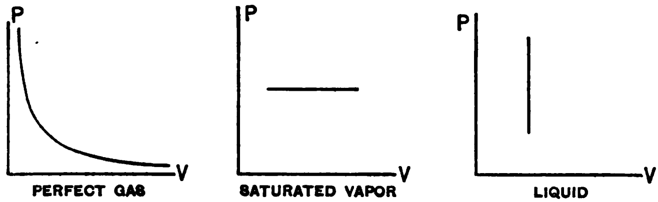


FIG. 141.

When a gas is at a temperature lower than the critical one, the process of complete liquefaction is in three stages: the gas is compressed at constant temperature until it becomes saturated, as it will at a definite pressure, depending upon the temperature; further decrease in volume produces condensation of the vapor in liquid form, but no change in pressure, since the temperature is constant; this condensation will continue until for a definite volume all the vapor has been liquefied; any further decrease in volume requires now a great increase in pressure. So the isothermal for the entire process is a combination of the three curves described above.

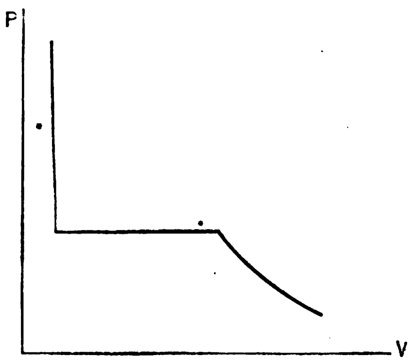
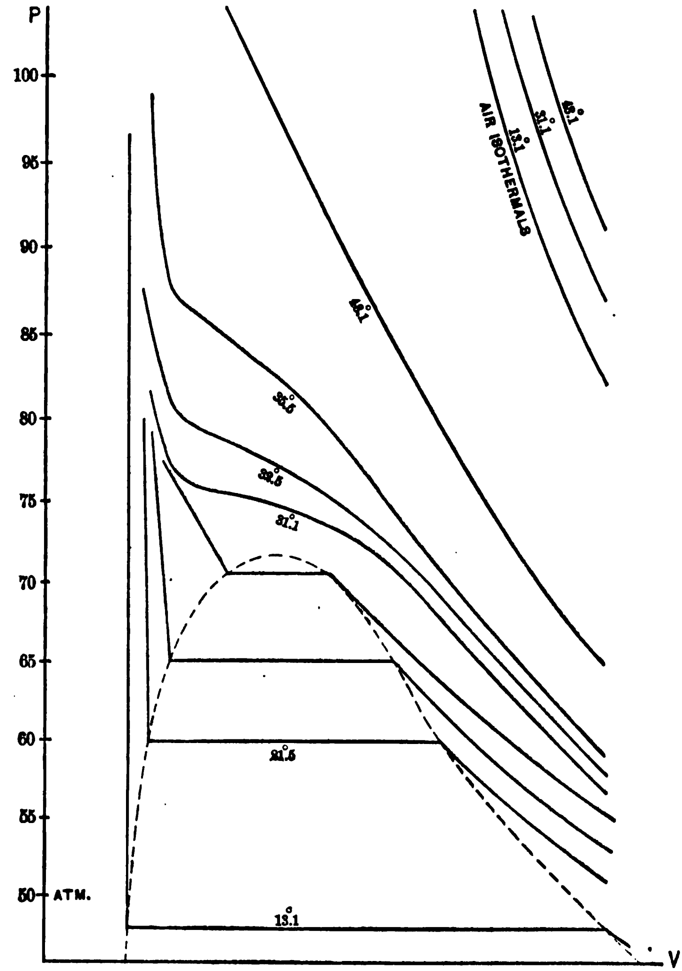


FIG. 142.



ISOTHERMALS OF CARBON DIOXIDE

FIG. 148.

For a different temperature there would be a different isothermal; and the accompanying figure shows a series of isothermals as actually observed for carbonic acid gas, CO_2. The critical temperature is that which corresponds

to the first isothermal which has no horizontal portion; because it is only when there is a horizontal portion that any liquefaction may be seen. The locus of the points which mark the right-hand ends of the horizontal lines, i. e. the points of saturated vapor, is called the "vapor curve." The locus of the left-hand ends, i. e. the points of complete condensation, is called the "liquid curve." These two curves meet in a point on the critical isothermal known as the "critical point." The pressure corresponding to this point is called the "critical pressure."

A liquid in a condition represented by a point on the diagram near the critical point is said to be in a "critical state," because any slight change may produce great disturbances, e. g. if the temperature is raised, the volume being kept constant, *all* the liquid instantly becomes vaporized.

Sublimation

201. Sublimation is a name given to the process by which a solid becomes a vapor without passing through the intermediate liquid form. Many solids do this. Snow, camphor, sulphur, and others often evaporate directly.

Certain laws have been observed. The most important is that, corresponding to any definite temperature, there is a certain pressure at which the solid and its vapor are in equilibrium, and conversely. The pressure which thus corresponds to a definite temperature may be measured by either the statical or the dynamical method. The latent heat of sublimation has been measured also.

Dissociation

202. Dissociation is the name given to that process by which the molecules of a substance are decomposed into simpler parts. It may be produced in various ways; but of course there are energy-changes, because the potential energy of a molecule is not the same as that of its dissociated parts.

One method of producing dissociation of a gas is to raise its temperature. Under these conditions many gases are partially dissociated. If the pressure of the mixture of gases, the original one and its parts, is observed, it is found that corresponding to any temperature there is a definite pressure at which the dissociation stops, i. e. a definite pressure of equilibrium; and the number of calories required to dissociate 1 gram at a given temperature is called the "heat of dissociation" at that temperature.

The amount of energy required to dissociate a compound, or the amount set free when the compound is formed (if such is the case), may be measured by ordinary calorimetric methods. Some compounds absorb energy when they are formed. The number of calories produced or required, if 1 gram of the compound is formed, is called the "heat of combination" of the substance.

There is good evidence for believing that, when some substances are dissolved in certain liquids, especially inorganic salts and acids in water, they are partially dissociated; but this effect is complicated by certain electrical phenomena to be discussed later.

Application of Principle of Stable Equilibrium. The effect of changing the temperature may be determined from the "principle of stable equilibrium." If the combination involves a development of heat-energy, then adding heat-energy will produce dissociation. If, on the contrary, combination is accompanied by an absorption of heat-energy, then taking away energy will produce dissociation or breaking up of the compound; and the addition of heat-energy must produce combination.

Solution

203. Whenever any two substances are placed in contact, solid, liquid, or gaseous, there is no reason for not thinking that each dissolves in the other somewhat; that is, that some molecules of each substance become

detached, as it were, and interpenetrate. This is known to occur in many cases; e. g. salt and water, alcohol and water, mercury and tin, etc. In all cases there are energy-changes, and so heat-effects, which can be measured. Sometimes energy is absorbed from surrounding bodies; at other times it is given out to them.

No process of solution proceeds indefinitely. For a given temperature there is a definite state of saturation. The effect of raising the temperature may be determined from the "principle of stable equilibrium." The saturated solution is in stable equilibrium. Imagine a further solution; and let the solution be one which involves a development of heat-energy. Hence, if heat-energy is added to this saturated solution, the solution must be resolved into its constituents; it becomes supersaturated. If, on the other hand, the solution was one which is accompanied by absorption of heat-energy, then adding heat-energy to the saturated solution would make the solution unsaturated, and more solution will take place if possible.

In some solutions, as noted above, there is evidence that the molecules which go into solution are dissociated. It will be shown later on that there is reason for believing that this is the case with all liquids which can conduct electricity. No pure liquid is a conductor, only certain solutions; and these solutions are probably dissociated molecules dissolved in the pure liquid.

CHAPTER V

TRANSFER OF HEAT-ENERGY

As explained in mechanics, there are two ways, in general, in which energy may be transferred from one point to another. One is by the energy being carried by a portion of matter; e. g. a cannon-ball conveys energy from the exploded powder in the cannon to the target or obstacle struck. The other is by waves; e. g. sound-waves convey energy from the sounding body to the ear.

There are illustrations of both these methods in the transfer of heat-energy, i. e. of energy from the molecules of one body to those of another. Three general processes are distinguished: Convection, Conduction, Radiation.

204. Convection. In this process some portion of the body whose temperature is raised moves away to other places in the body where the temperature is lower, thus tending to raise the temperature throughout. Portions of solids cannot move about; so convection is limited to fluids, i. e. to liquids and gases. If a vessel of any liquid is placed over a fire, the portions of the liquid near the bottom become hotter and so expand; but on expansion the portions become less dense, and so tend to move upward towards the top of the liquid. Thus the energy is transferred. If the liquid was heated on top, there would be no convection. Similarly, there can be convection in a gas, if it is heated from below.

The process of convection is of fundamental importance in draughts of chimneys, ventilation, and in the great phenomena of nature, such as winds and ocean-currents.

205. Conduction. In conduction the separate portions of matter do not move bodily and thus transfer the energy; but they hand it on from one portion to the other. Thus, if a poker has one end in a fire, that end has its temperature raised; and this in turn raises the temperature of the portion next it, and so on down the entire length of the poker. The other end finally has its temperature raised, but not until all the intervening portions have had their temperatures also raised. It is found by experiment that all solids which can conduct electricity, i. e. all metals, are good heat-conductors; other solids are not. Of the liquids mercury is the only good conductor; and all gases are poor conductors. (To study conduction in fluids, the heat-energy must, of course, be supplied from above, so as to avoid convection.)

There are many illustrations of the power of metals to conduct away heat-energy. One of the most interesting is the ordinary miner's safety-lamp, in which the flame is surrounded by metal-gauze and solid metal pieces, so that the heat-energy from the flame is conducted away by the metal; and thus the temperature of the escaping gases is too low to ignite outside the lamp.

The following table gives an approximate idea of the conducting power of different bodies: —

TABLE XI

THERMAL CONDUCTIVITIES

Copper . . .	.96	Glass . . .	.0005
Iron	.20	Wool . . .	.00012
Stone . . .	.006	Paper . . .	.000094
Water . . .	.002	Air	.000049

It should be noted that the conducting power of any substance is different for different temperatures.

206. Radiation. This is the process by which the heat-energy is transferred from one point to another by means of waves. The existence of these waves may be demonstrated by the usual tests of wave-motion, which will be explained later under the section LIGHT; they advance with a finite velocity; they can interfere; they can be diffracted. Further, they can be polarized; and this, as we shall see, is proof that the waves are transverse. These waves, too, are not in ordinary matter, as are sound-waves, but can be propagated across a vacuum, or from the sun and stars to the earth. Since waves must be in some medium, this proves that there is a medium which permeates all spaces, large and small, between portions of ordinary matter. This medium, whose existence is thus demonstrated, is called "the ether." It must have inertia, because a finite time is taken for the propagation of the waves across a finite space. But the other properties of the ether are quite unknown.

There are, then, waves in the ether which carry energy away from various sources; and their properties are now well known. They are transverse; they travel through the ether with a definite velocity which is independent of the wave-length; but when portions of ordinary matter are immersed in the ether the velocity of the waves is changed, and is changed differently for different waves; the wave-lengths may be measured by suitable means; and their intensity may be determined. When these waves reach any obstacle, there will be reflected and refracted waves in general, if the obstacle is large compared with the length of the waves; and the laws of the reflected and refracted waves will be discussed later in LIGHT. It is sufficient to say here that all these waves obey the same laws as so-called light-waves. (See below.) Some of the energy will also be given the obstacle, i. e. will be "absorbed" by it. The effect of this absorption on the obstacle depends upon what work is done there as a result of the energy being

received from the waves. It is found by experiment that, if the obstacle is any ordinary body, the absorbed energy in general produces heat-effects. But, if the waves have lengths within certain limits, and if they fall on the eye, some of the energy is used in producing the sensation Light. Again, if the obstacle is one of many chemical substances, e. g. a photographic film, some of the energy goes to causing certain chemical reactions. If the waves are longer than any of these, their energy may be consumed in producing electrical oscillations in conductors.

Ether-waves, then, produce heat-effects when they are absorbed, only if they are absorbed by special bodies. Since in some cases they produce the sensation light, this fact renders easy the study of their laws. As will be proved later, it is known that the velocity of these waves through the pure ether is almost exactly 3×10^{10} cm. per sec. Further, those waves which produce light-sensation have wave-lengths lying roughly between 0.00004 and 0.00008 cm., i. e. are about $\frac{1}{45000}$th of an inch long. Waves as short as 0.00001 cm. and as long as 0.002 cm. have been measured. The energy carried by the waves, or their intensity, may best be studied by letting waves of different lengths be absorbed by some body, such that all the energy goes into heat-effects. By suitable means the entire series of trains of waves emitted from any source may be analyzed into separate trains of waves, each train having a definite wave-length which differs minutely from those of its neighbors. Such an analysis may be done by a prism or a grating (see LIGHT). If, then, a curve is drawn, the locus of points whose horizontal distances from some fixed line give wave-lengths, and whose vertical distances from another fixed line perpendicular to the first give intensities, it is called the "energy-curve" for the particular source. Certain interesting laws have been observed between energy-curves and temperatures of the source. Thus, the higher the temperature of the source, so much further does the energy-

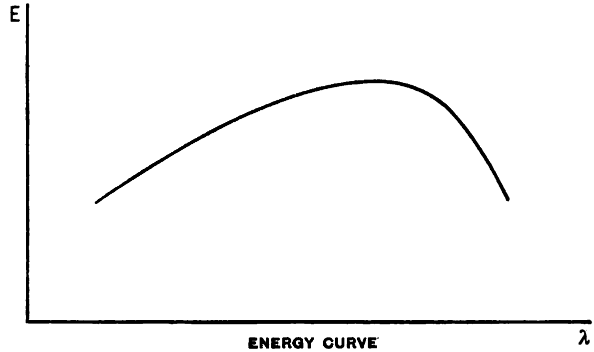


Fig. 144.

curve extend towards the left; that is, there are now, among the waves emitted, some shorter than there were before. At the same time, as the temperature is raised, the intensity of each particular train of waves is increased; or the energy-curve is raised, — not uniformly, however.

The energy in these ether-waves is emitted by different forms of matter. In fact every piece of matter in the universe is undoubtedly producing ether-waves. A solid or a liquid is sending out what is called a "continuous spectrum;" that is, in the series of waves emitted, there are trains of all possible wave-lengths in between certain limits, there are no special waves absent. A gas or vapor, on the other hand, produces a "discontinuous" spectrum; certain waves are present, others are absent. The ones which are present depend upon the condition and nature of the gas or vapor.

Few bodies absorb practically all the waves which fall upon them; most of them absorb only certain definite waves. And, further, the waves which are absorbed and the intensity of the absorption depend upon the temperature of the absorbing body. There must be some connection between the radiating power of a body and its absorptive power. For, consider a body enclosed in a space,

e. g. a hollow sphere, whose bounding surface is kept at a constant temperature, and which is lined with some substance having the power to emit all wave-lengths up to a certain limit. Some of these waves are reflected from the body, some are transmitted, some are absorbed. But some waves are also emitted by the body itself; and, since experiments prove that the temperature of the body inserted gradually becomes equal to that of the enclosing walls and then does not change, the waves emitted when this temperature is reached must be identical in every respect with those absorbed. In other words, the emissive power of any body at any temperature exactly equals the absorptive power at that same temperature. Thus, red glass is called so, because it absorbs all the colors which come from any hot source, except red. But if the glass is raised to that temperature, it will emit those waves which it absorbed before, viz. bluish green, as will be apparent if it is carried while hot into a dark room. Sodium vapor at a high temperature has the power of emitting a few characteristic waves, especially two which produce the sensation yellow. So, if white light, i. e. waves from a very hot solid, is passed through sodium vapor at a low temperature, these particular waves are absorbed and are therefore absent in the transmitted waves. This is, of course, a general mechanical principle, which has been illustrated before in the case of two tuning-forks (see Art. 129). If a fork is set in vibration, the waves produced falling upon another fork of the *same* frequency will set it in vibration. That is, the second fork absorbs the energy from those waves which it itself could produce, because it has the same frequency.

207. Reflection, Emission, Absorption, Transmission. A body which absorbs well can also emit energy well; as is shown in all blackened bodies which are not polished. But, when energy, in the form of waves, falls upon a body, part is reflected, and the rest absorbed or transmitted; so

that, if a body reflects well, it must absorb poorly. Thus, polished surfaces, particularly metals, must absorb very little energy, and so must emit very little, because they are such good reflectors.

Again, the amount of energy which is transmitted depends upon the amount reflected and absorbed. An equation may be written: Incident energy = reflected + absorbed + transmitted.

The question of the transparency of any body to these ether-waves depends largely on their wave-length (or, better, their wave-number) and their intensity. Many substances, transparent to visible waves, are almost entirely opaque to longer or shorter waves. Thus, glass transmits freely light-waves, but shuts off the others. A solution of iodine in carbon bisulphide is entirely opaque to light-waves, but is transparent to the longer waves. And in every case, the more intense the waves, i. e. the higher the temperature of the source, so much the greater are the chances of any wave getting through any substance.

208. Flow of Heat-Energy. In Convection and Conduction it is obvious that heat-energy is carried from places of high temperature to places of low; and the rapidity of this flow varies directly as the difference in temperature of the two places.

The same statements are also true of Radiation. If a body at a low temperature is placed inside a second body which is so made as to emit no waves from the outer surface, and which is at a higher temperature than the first body, each radiates a certain amount of energy depending only on its own temperature and condition; each body also absorbs a certain amount of the energy radiated by the other. The hot body, though, radiates more energy than it absorbs; while the opposite is true of the cold body. Consequently, the temperature of the hot body falls, and that of the cold body rises, until the temperatures of the two are the same.

Thus, heat-energy has passed from a body at a high temperature to a body at a lower one; and the rapidity of the process must vary directly as the difference in temperature.

If there is no difference of temperature, there is no flow of heat-energy; and, if there is a difference, the flow is always from high to low temperature.

The analogy should be noticed to the flow of a fluid from places of high pressure to places of low. For instance, water flows from a high tank to a lower one; a gas compressed into a vessel will expand, if it can, into another vessel where the pressure is less. And the rate of the flow varies directly as the difference in pressure. At a great pressure the fluid has more energy than at a lower one; so, in the flow of the fluid, what may be called "volume-energy" flows from high to low pressure.

CHAPTER VI

KINETIC THEORY OF MATTER

209. IT is impossible to understand how waves can be produced in the ether by means of ordinary matter, or how the waves can convey heat-energy from one portion of matter to another, unless there is some connection between matter and the ether, and unless the smallest portions of matter are in motion. Waves in the ether must therefore be considered as produced by the vibrations of extremely small portions of matter. Consequently all portions of matter, however small, are considered as being in motion; the molecules are thought to be moving bodily; and the parts of the molecules, if there are any, may be making vibrations.

210. States of Matter: Solids. The idea of a solid is, then, a fixed configuration of molecules, where the molecules may vibrate about positions of equilibrium but cannot move from one point to another in the solid. The reason why a solid emits a continuous spectrum is because the vibrations of the portions of the molecules are so hampered by the close proximity of other molecules that they vibrate in all periods within certain limits, and so produce waves of all lengths. The same explanation applies to the waves emitted by a liquid.

Liquids. In a liquid the molecules have increased freedom of motion, and can move about from one point to another. They undoubtedly are moving with considerable velocity; and so it is easy to understand how some may escape from the surface and thus cause evaporation.

Vapors. In a vapor or gas the molecules are no longer close together; and they are moving with great velocity in all directions. In the vapor over a liquid, some of the molecules in their motion will undoubtedly strike the surface of the liquid, and become entangled. There is, thus, a process of continuous evaporation and condensation; and the state of saturated vapor is reached when the amount evaporated in a definite time equals that condensed in the same time. Similarly, in dissociation there is undoubtedly a state of continuous dissociation and combination existing; and equilibrium is reached when the two opposite processes are equal.

Gases. In a gas the molecules move freely in the intervals of time between collisions with each other or with the walls of the vessel; and so they have what is called "a free path." During their passage along their free paths, the molecules are uninfluenced by each other; and so the portions of the molecules may by their vibrations send out waves of certain definite wave-numbers. It is easy, too, to understand conduction of heat-energy, diffusion, and other properties of gases on this kinetic theory.

FIG. 145.

If a closed space is so thoroughly exhausted by means of air-pumps as to allow the molecules of a gas to have free paths of several inches, the entire nature of the gas changes, and many new properties are observed. This condition is sometimes called a fourth state of matter. One of the most interesting properties of this state is that there are curious motions produced in solid bodies placed in the exhausted space, when ether-waves fall upon them. If a vane, made of two arms carrying at each end a plate of mica polished on one side

and blackened on the other, is balanced on an axis in a glass bulb which is thus exhausted, it will revolve when any hot object is brought near, the blackened sides of the mica always moving away from the object. The explanation is not difficult. The blackened surface absorbs the energy from the ether-waves; its temperature is raised; and so the molecules, as they rebound from the surface, receive an extra velocity. Owing to the reaction of the molecules on the mica surface, the latter is driven back, away from the hot body emitting the waves. This phenomenon would not take place in an ordinary gas, because any rise in temperature or increased velocity of some of the molecules would be almost instantly communicated to the rest of the gas; and so there would be an equal force on both sides of the mica-vanes.

211. Kinetic Theory of Gases. There is a definite theory of gases, more specific than the general one given above; and it states that it is possible to prove that a collection of small, hard, smooth, elastic spheres, identically alike, thrown at random into a large space with rigid walls will have certain properties identical with those of perfect gases. It is necessary, further, to assume that the space is so large and the spheres so numerous that the "principle of statistics" may be applied. That is, that, although the properties of an individual sphere may change each instant, the average for all the spheres of the same property (e. g. speed) will not change unless the conditions are changed.

a. *Energy and Temperature.* A set of spheres like the one described consists of a great number, each one having the same mass, and each one moving with a definite velocity at any instant. The velocities of the different spheres vary widely; and that of any individual sphere changes at each collision. But at any instant there is a certain *average* velocity in any direction; and there is also a certain average kinetic energy. Since the particles are perfectly

smooth and elastic, the entire kinetic energy of translation is not changed by any collisions; and so, if the walls of the vessel are rigid, the average kinetic energy cannot change. But, if the walls are free to move out, they will do so under the pressure produced by the impacts of the spheres, external work will be done; and so the average kinetic energy of translation will decrease. Conversely, if the walls are forced in by any external force, the average kinetic energy of translation is increased. These properties of the average kinetic energy of translation of the set of spheres correspond perfectly to those of the temperature of an ordinary gas. Further, it may be proved mathematically that two sets of spheres, thoroughly mixed, will be in equilibrium when the average kinetic energies of translation of the two sets are equal. That is: if m_1 is the mass of any one sphere of the one set, and u_1 is the average velocity of that set, while m_2 and u_2 correspond to the second set, there will be equilibrium when

$$m_1 u_1^2 = m_2 u_2^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

And two gases are in equilibrium if their temperatures are the same; so, on the kinetic theory, this equation may be considered as the mathematical condition for equality of temperature of two gases.

b. *Momentum and Pressure.* As the spheres strike the walls of the vessel, their velocity perpendicular to the walls is reversed, and so there is a change in momentum. But, whenever there is a change in momentum, there is a force produced; so the spheres exert a force on the walls, normally outward. It is not difficult to calculate this force. Consider a cubical vessel containing the set of spheres; and let each edge have a length a. Since the motion of the spheres is entirely a random one, there is as much momentum parallel to one edge of the cube as to any other. So the entire set of spheres may be considered as divided into three equal sets, each mov-

ing parallel to one edge with a certain average velocity. Thus, let

m = mass of each sphere,
n = number of spheres in 1 cubic centimetre,
u = average velocity parallel to one edge.

Hence there are $n\,a^3$ spheres in the cube, and $\frac{1}{3}\,n\,a^3$ are moving parallel to each edge. These strike the end side with a momentum $\frac{1}{3}\,n\,a^3 \,.\, m\,u$; but this is reversed at the wall, that is, the velocity becomes $-u$; and so the *change* in the momentum is $\frac{2}{3}\,n\,a^3 \,.\, m\,u$. It takes each sphere, moving with the velocity u, a time $\frac{2\,a}{u}$ for it to pass over to the opposite wall and back again; or, in other words, each sphere returns and collides with the same wall $\frac{u}{2\,a}$ times in one second. Consequently the entire change of momentum at the wall in one second is

$$\tfrac{2}{3}\,n\,a^3\,m\,u \cdot \frac{u}{2\,a} = \tfrac{1}{3}\,m\,n\,u^2\,a^2.$$

By the definition of force (Art. 30) this is the force acting on the wall whose area is a^2. Hence the pressure, the force per unit area, is

$$p = \tfrac{1}{3}\,m\,n\,u^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

m is the mass of each sphere, n is the number of spheres in 1 cc.; hence $m\,n$ is the mass in 1 cc., i. e. the density, ρ. So

$$p = \tfrac{1}{3}\,u^2\,\rho \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2\,a)$$

But since, if the average kinetic energy of translation is constant, u^2 is also, this formula may be read: If the average kinetic energy of translation of a set of spheres is constant, the pressure on the walls is proportional to the density. This is identical with Boyle's laws for perfect gases.

Further, assuming this law to be rigidly true for actual gases, the average velocity of a molecule of a gas may be calculated for any pressure. Substitute for p its value, and for ρ its value for the particular value of p; and then $u^2 = 3p/\rho$.

If there are several sets of spheres enclosed in the same space, each will obviously produce its own pressure on the walls of the vessel, provided that the space is large enough and that one set does not vastly outnumber any other. This is Dalton's Law. Thus, the entire pressure

$$P = p_1 + p_2 + p_3 + \text{etc.},$$

where
$$p_1 = \tfrac{1}{3} m_1 n_1 u_1^2,$$
$$p_2 = \tfrac{1}{3} m_2 n_2 u_2^2, \qquad \text{etc.}$$

If there is equilibrium of the sets of spheres, equation (1) requires that

$$m_1 u_1^2 = m_2 u_2^2 = m_3 u_3^2 = \text{etc.} \equiv c \text{ (say)}.$$

Hence
$$P = \tfrac{1}{3} c (n_1 + n_2 + n_3 + \text{etc.}).$$

That is, the pressure in a mixture of gases, at a given temperature depends simply upon the total *number* of particles, not in the least upon the mass of any individual component. (This statement may be regarded as an extension of Avogadro's Law, which is given below.)

c. *Expansion.* It may be shown, too, that Charles' Law is true for sets of spheres. That is, if the volume is kept constant, but the kinetic energy increased, the pressure will increase at a rate which is the same for all sets of spheres.

d. *Avogadro's Law.* If there are two sets of spheres, at the same temperature and pressure, equations (1) and (2) give

$$m_1 u_1^2 = m_2 u_2^2$$
$$m_1 n_1 u_1^2 = m_2 n_2 u_2^2.$$

And so
$$n_1 = n_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

or, the number of spheres of the first set in 1 cc. is the same as the number of spheres of the second set in an equal volume.

Applied to gases, this principle is called Avogadro's Law; and it states that, if two gases are under the same conditions of temperature and pressure, they both have the same number of molecules (or smallest parts) in equal volumes.

The density of a gas, ρ, equals $m\,n$; and so the ratio of the densities of two gases at the same temperature and pressure is

$$\rho_1 / \rho_2 = m_1 n / m_2 n = m_1 / m_2 \quad . \quad . \quad . \quad (4)$$

m_1 is the actual mass of a single molecule; and this formula gives a method for comparing the masses of molecules of different gases, because the densities are measurable quantities.

What is called the "molecular weight" of a gas in Chemistry is a number which is proportional to the mass of a single molecule of the gas, the factor of proportionality being so chosen that the molecular weight of hydrogen is 2. So the molecular weights of two gases, w_1 and w_2, must be in the same ratio as m_1 and m_2; and there will be the same number, N, of molecules of any gas in a mass of that gas equal to its molecular weight.

Boyle's and Charles' laws for a gas give

$$\frac{p\,v}{T} = R\,M,$$

where M is the mass occupying a volume v. Let this mass be w, that is, a number of grams which equals the molecular weight; there are then N molecules in the space v. Consider a second gas whose equation is

$$\frac{p'\,v'}{T'} = R'\,M',$$

and take w' grams, where w' is its molecular weight. There are N molecules of this second gas in the volume v'. But equal numbers of molecules of different gases occupy equal volumes if temperature and pressure are the same. Hence, when

$$M = w, M' = w', \quad v' = v \text{ if } p = p' \text{ and } T = T'.$$

Or

$$R\,w = R'\,w' = R_0, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where R_0 is a constant for all gases. So, if $M = q\,w$, that is, if the number of grams of the gas is equal to q times the molecular weight, the general law for a gas becomes

$$\frac{p\,v}{T} = q\,R\,w = q\,R_0.$$

Writing $v / q = v_0$, or the volume which would be occupied by a mass of the gas equal to the molecular weight, this becomes

$$\frac{p\,v_0}{T} = R_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Avogadro's law (or hypothesis) is the basis of many chemical theories; and, although not rigidly proved for actual gases, it must be reasonably correct. It is observed that, when equal volumes of hydrogen and chlorine gas combine to form hydrochloric acid gas, the volume formed is the sum of the two volumes of the constituents, if the pressure and temperature are kept constant. The chemical formula, as ordinarily written, is

$$\mathrm{H} + \mathrm{Cl} = \mathrm{HCl}.$$

On its face, this would mean that one molecule was formed from two; but this, on Avogadro's hypothesis, would require that the resulting volume of HCl should be one-half that of the sum of the two component volumes, if pressure and temperature are the same. This discrepancy between hypothesis and experiment may be explained if it

is assumed that a molecule of H has two atoms, and one of Cl has two atoms; then HCl is a compound of an atom of H and an atom of Cl. Thus, if there are *n* molecules of H in a volume *v*, there are also *n* molecules of Cl in an equal volume, if temperature and pressure are the same. On combination, as expressed in the formula

$$H_2 + Cl_2 = 2\,HCl,$$

2 *n* molecules HCl are formed; and so twice the volume is necessary to keep the same temperature and pressure.

Similarly, when water-vapor is formed, 2 *n* molecules of H combine with *n* molecules of O to form 2 *n* molecules H_2O. So the steam must be compressed into a volume which is twice the volume *v*, if the pressure and temperature are to be kept unchanged. That is, there is a "condensation" from volume 3 *v* (2 *v* for H and *v* for O) to volume 2 *v*.

In general, when any compound gas is formed from other gases, the volumes bear simple mathematical relations to each other, if the temperature and pressure are the same for all the gases; and this leads, on Avogadro's hypothesis, to a definite idea of the nature of the molecule.

CHAPTER VII

THERMODYNAMICS

THERMODYNAMICS is the name given to the science which studies the application of mechanical principles and equations to the physical properties of heat-energy.

212. First Principle of Thermodynamics. This is nothing but the application of the principle of the conservation of energy to heat-effects. It asserts that, whenever energy is spent in producing heat-effects, the amount of work done exactly equals the sum of the internal and the external work (see Art. 167). This is perfectly in accord with all experiments. The amount of work required to produce any "effect of heat" is the same, no matter how the work is done. Thus, the amount of energy necessary to raise the temperature of 1 gram of water from 10° to 11° C. is the same if the work is done by turning a paddle in the water, or if the energy is produced by heating a wire by means of an electric current. This amount of work, called J, the "mechanical equivalent of Heat," has, as has been said, the value 4.2×10^7. Further, J, as found by means of the equation

$$J(C_p - C_v) = R$$

(see Art. 182) is the same, within the limits of accuracy of the experimental determination of C_p, C_v, and R.

213. Second Principle of Thermodynamics. This principle asserts what must be regarded as an axiom, viz. that heat-energy of itself cannot pass from one body to another if the first body is at a lower temperature than the second.

There are many most important consequences of this principle; but they are not suited for presentation here.

A heat-engine is a mechanism which allows some "working-substance," like steam, to receive heat-energy at a high temperature; allows the working-substance to do external work, in doing which its temperature falls; then by means of external work brings the working-substance back to its original condition; etc. indefinitely. While the working-substance is being brought back to its original condition, it must give out some heat-energy to some external body. Let h_1 be the heat-energy received, h_2 be the heat-energy given out. Then, by the conservation of energy, the external work done by the substance in excess of that done on it is

$$W = h_1 - h_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

because the working-substance is back in its original condition; and so no internal work has been done, only external.

The "efficiency" of a heat-engine is, by definition, the ratio W/h_1; and it may be proved by the second principle of thermodynamics that this efficiency cannot, under definite conditions of temperature, exceed a certain limit. An engine which has the greatest possible efficiency is called a perfect engine, and it may be proved that the efficiency of such an engine is (except for a small number of bodies) entirely independent of the working-substance, being the same for all fluids and also certain solids. But the efficiency does depend upon the conditions of temperature. If the body from which the heat-energy is taken into the working-substance has an absolute temperature, T_1 ($t°$ C. + 273), and if the absolute temperature of the coldest large body available for the working-substance to give up heat-energy to is T_2, it may be proved that the efficiency of a *perfect gas-engine* is $\dfrac{T_1 - T_2}{T_1}$. But the efficiency of any other perfect

engine working between the same temperatures must be the same. (This leads to a numerical determination of temperature, which is entirely independent of the nature of the thermometer.)

The greatest possible efficiency is, of course, unity; i. e. under these conditions,

$$\frac{T_1 - T_2}{T_1} = 1.$$

Hence T_2 must equal 0. A lower temperature than this would mean an efficiency greater than one, which is impossible. So the lowest possible temperature is that for which $T = 0$ or $-273°$ centigrade. This is called the "absolute zero."

BOOK IV

ELECTRICITY AND MAGNETISM

BOOK IV

ELECTRICITY AND MAGNETISM

INTRODUCTION

IT has been known for many hundred, if not thousand, years that, when a piece of amber is rubbed with flannel, it has the power of "attracting" small portions of matter, and that a magnetic needle suspended on a pivot turns so as to point in a northerly direction. Again, no phenomenon has been observed with such unfailing interest as the discharge of lightning. These phenomena are all related to each other; and the study of their laws and those of other phenomena of the same nature forms the subject of "Electricity and Magnetism."

214. All electric and magnetic effects are intimately associated with the "ether," that medium which conveys the waves discussed in Radiation, Chapter V. of HEAT; and it may be well to describe briefly its properties. It is a medium which penetrates all the spaces, large and small, between the portions of ordinary matter; in fact, the ether may be regarded as a universal medium in which the minute portions of ordinary matter are immersed. The ether has inertia, as is proved by the fact that waves in it travel with a finite velocity; but its other properties are uncertain. If it is made up of parts, as ordinary matter is, they must be extremely minute; because waves in any medium are always immensely longer than the dimensions of the ulti-

mate particles of that medium, and ether-waves are themselves very short. Ether-waves are, however, transverse (as will be shown later); and this means that the ether has rigidity, or resists a change in shape (see Art. 74). So the ether may have a strain of this type; and, if it does, it will have potential energy. The ether can also move; and then it will have kinetic energy. If there is any ordinary matter immersed in the ether, it is connected to the ether in some way; so that all the properties of the ether are affected by the presence of the matter, it is "loaded down" with it. If the ether is strained, so also are any portions of matter in it; and the apparent inertia of the ether is evidently increased by the presence of ordinary matter. Consequently, in all electric and magnetic phenomena the effect of matter is most important. It would be expected that different kinds of matter would produce different effects; and that, in the case of waves, those of different length would not be affected alike. Such is observed to be the case.

CHAPTER I

GENERAL PROPERTIES OF ELECTRIC CHARGES

215. Production of Charges. If two different bodies, e. g. glass and silk, or sealing-wax and flannel, are rubbed together so as to bring their surfaces into intimate contact, and are then separated, two things may be observed: (1) Work is required to separate them; and so potential energy is stored up somewhere, in the form of a strain. This is proved by the fact that, if, after separation to a short distance, the two bodies are set free to move, they will move towards each other. (Compare raising a body off the earth, the stretching of an elastic cord, etc.) (2) Each body now has the power of causing small bits of matter, like dust, pieces of paper, pieces of gold-foil, etc., to move towards it. (This property will be shown in Article 221 to be due to changes in the strain of the medium around the charged body.)

This last property is generally taken as the test of "electrification." A body which has it is said to be "electrified" or "charged," or to have a "charge" of electricity. If two bodies which differ in the least in the nature or arrangement of their molecules are brought into contact in any way, e. g. by pressure or by friction, they are found, on separation, to be charged.

There are many other methods of producing charges; and some will be discussed in later chapters. One of the commonest is simply to touch a charged body with an uncharged one, when some of the charge passes to the latter.

The fact that the medium around a charged body is strained is proved also by the disruption of air, glass, etc. by ordinary sparks from electric machines. The strain primarily is in the ether, as is proved by the possibility of having charged bodies in a vacuum; but, if ordinary matter is present, it as well as the ether is the seat of the strain (except in conductors. See Art. 220). It will become evident in the following paragraphs that the essential feature of electrification is the particular strain which is characteristic of it.

216. Positive and Negative Charges. Although all charges can produce motions in small portions of matter, there are different kinds of charges. This may be shown by the following experiments: Suspend in a paper stirrup carried by a long thread a glass rod which has been rubbed with silk; bring near its end another glass rod which has also been rubbed with silk; there is *repulsion*. If a piece of sealing-wax (or ebonite) which has been rubbed with flannel is brought near the suspended glass rod, there is *attraction*. If the charged sealing-wax is suspended in the stirrup instead of the charged glass rod, it may be observed that a glass rod charged by rubbing with silk will produce attraction, while another rod of sealing-wax charged by rubbing with flannel will produce repulsion.

Thus, two identical bodies charged by the same process cause repulsion. It is also found, by experiments performed on the suspended glass rod and the suspended sealing-wax, that all other charged bodies may be divided into two classes, — some attract the charged glass, others repel it; but those which attract the glass repel the charged sealing-wax, and those which repel the glass attract the sealing-wax. Glass rubbed with silk belongs to the second class; sealing-wax rubbed with flannel, to the first. And the general law may be stated, that "all charges of the same class repel each other, and attract charges of the other class."

The same material may belong to both classes, depending upon what other body it has been in contact with. Thus, glass rubbed with the fur of a cat's skin *attracts* glass rubbed with silk. So the above division is one of charges, not of materials.

These two classes of charges have received the names "positive" and "negative." One reason for this choice is that, if one kind of a charge produces any kind of motion in a charged body, the other kind of a charge will produce exactly the opposite motion of the same body; just as a positive force or moment produces exactly opposite effects to those of a negative force or moment.

By common consent, the kind of charge which a glass rod takes when rubbed with silk is called "positive;" and so the nature of any charge may be at once determined by comparing it with that on the glass.

217. "Specific Attraction" of Matter for Electricity. When two different bodies are brought into contact, and then separated, attraction is observed between the two; that is, one is charged positively, the other negatively. The question as to which will be positive and which negative depends upon the properties of the two bodies with relation to each other.

When any one body becomes charged, there is a change in the potential energy of the molecules of the body; but this change may be different for a positive charge and for a negative one. Therefore, since it is a general law of nature that potential energy tends to get as small as possible, this particular body will have a tendency to become charged with that charge, + or −, which will produce the greater decrease in the potential energy. This is sometimes expressed by saying that each kind of matter has a "specific attraction" of a definite amount for either positive or negative electricity.

If, then, two bodies are brought in contact and charges are produced, that one which has the greater specific at-

traction for a positive charge will become positively charged, and the other negatively. (See the following article.)

It is possible, of course, to arrange a table of substances in which they will stand in the order of their specific attractions for positive charges.

218. Quantities of Electricity. Since charges produce certain motions (or "forces") in other charged bodies, the equality of two charges may be tested by seeing if they will produce identically the same effect on a third charged body under the same conditions. If they do, they are said to have equal "quantities" of electricities. If, however, one charge produces a force which is exactly the opposite of that produced by the second one, the two charges have equal quantities, but are of opposite kinds. One has a charge $+ e$; the other, a charge $- e$.

If two bodies are brought in contact and separated, it may be observed that one is charged positively, and the other negatively; and it may be proved by most careful experiments that the quantity of positive charge of the one exactly equals that of the negative charge of the other. In other words, a positive charge can never be obtained without at the same time there being an equal negative charge produced. This may be regarded as a necessary consequence of the fact that the medium around charges is strained. For a strain can be maintained only by two *equal* and *opposite* forces; e. g. a wire stretched or twisted, a liquid compressed, etc. So in an electric strain there must be equal and opposite phenomena at the ends of the strain; that is, where it starts and ends.

In order to measure quantities of electricity, that is, in order to give a numerical value to a charge, it is necessary to adopt some unit charge to serve as a standard. The unit adopted for certain measurements is defined in this way: two unit charges placed at a distance of one centimetre apart in air have a force between them of one dyne. If one charge is positive and the other negative, the force

will be one of attraction; otherwise there will be repulsion. This force can, of course, be measured (at least theoretically) by means of a spiral spring or a chemical balance. The unit thus defined is called the "Electrostatic Unit of Quantity," because it is used in measuring charges which are at rest.

219. Law of Electrostatic Force. The force between two charged bodies (i. e. the change in momentum in one second which would be produced in each body if free to move) is found to depend upon three things: the quantities of the charges, their distance apart, and the material medium in which the charged bodies are placed, e. g. air, water, paraffin, etc. The exact law of the action may be expressed,

$$F = \frac{e e'}{K r^2}, \quad \ldots \ldots \ldots \quad (1)$$

where e and e' are the quantities of the two charges; r is their distance apart; and K is a constant for any one medium, but different for different media. F is the mechanical force, and is measured in dynes; it is positive for a repulsion, because, if e and e' both have the same sign, i. e. if there is repulsion, F is evidently positive.

On the electrostatic system of units, K has the value 1 for air, for a unit charge is so defined that, if placed *in air* at a distance of 1 cm. from an equal charge, there is a force of 1 dyne. That is, if $e = e' = 1$ and $r = 1$ in air, $F = 1$. Hence $K = 1$ for air. For other media K has a value always (except in certain gases at low pressure) greater than for air. On any other system of units than the electrostatic one, K would not equal 1 for air. K varies with the temperature, and also the pressure, if the medium is a gas.

This law of force is called "Coulomb's Law," because he first verified it experimentally. His method was the direct one of placing charged bodies at certain distances apart and comparing the forces. A better method is to make

deductions from the law, and see if they are all fulfilled. Such has been done, and it may be proved that many phenomena are easily explained as consequences of this law, and cannot be accounted for by any other law.

220. Conductors and Dielectrics. It is easily proved by experiment that, so far as electric charges are concerned, there are two classes of material bodies. If a charge is placed on a body of one class, it stays at the point where it is placed; e. g. if a certain portion of a glass rod is rubbed with silk, that particular portion is the only part of the glass which receives a charge. Such bodies are called "insulators" or "non-conductors."

But, if a charge is placed on a body of the other class, it spreads and appears over its entire outer surface. Such a body is called a "conductor;" and all metals are illustrations. The fact that the charge on a conductor is entirely on the outer surface is most important. It can be proved by delicate experiments. If the conductor is hollow, the entire charge is on the outer surface, as stated, unless another charged body is placed inside, in the hollow space, in which case there will be a charge on the inner surface of the conductor caused by the presence, in the interior, of the second charged body. (See Art. 232.)

If two conductors are charged, it is possible to keep the charges unchanged, only if the two conductors are separated by means of an insulating substance. For, if a conductor is used, the charges spread over it, and so everything is changed. If the original charges were e and e', the entire charge when the two conductors are joined by a third one is found by experiment to be in every case $e + e'$. So, if $e' = -e$, that is, if the two charges were equal and opposite, the final charge is zero, the two charges have neutralized each other. Any uncharged conductor may therefore be considered as having on its surface two equal and opposite charges of any amount desired, and it is possible, as will be shown later (Art. 229), to separate these charges.

The following table contains the names of some non-conductors and conductors : —

TABLE

Non-Conductors.	Conductors.
Glass.	All metals.
Paraffin.	Salt water.
Dry air.	Moist cotton.
Ebonite.	The human body.
Silk.	Damp wood.
Porcelain.	
Shellac.	
Wool.	
Resin.	
Oils.	

The facts that a charge spreads itself over the surface of a conductor, and that consequently, in order to keep a charge on a conductor, it must be surrounded by a non-conductor, may be explained, if it is remembered that the essential feature of a charge is the strain in the surrounding medium. If glass and silk are rubbed together and then separated, the medium in between them and around them is strained, as already explained. The electric strain in the medium begins and ends on the charges, just as, in the mechanical strain produced in a wire by stretching it, the strain is bounded by the two ends where the equal and opposite forces are applied. In fact, the electric charges may best be regarded as simply the phenomena at the ends of the strain. To be strained, the medium must have some elasticity; but imagine a medium which has none, which offers no opposition to any attempt to strain it. It would be impossible to keep a positive and negative charge apart by such a medium. Further, imagine a charge placed at any point in the air; the air will be strained on all sides of the charge. Now place this charge on a small portion of a medium like the one described, which offers no resist-

ance to be strained; there can be no strain, of course, in this medium, and so the strain in the surrounding air must begin at the outer *surface* of the inner medium. Consequently, the action is just as if the charge had been spread over the surface of that medium. So, a conductor behaves exactly like a medium which cannot resist any attempt to strain it, which yields at once to the slightest electric stress. A non-conductor, on the other hand, can be electrically strained; and, since the essential features of charges are in the strained medium, a special name has been given non-conductors so as to emphasize this fact: they are called "dielectrics."

In the law of electric force $\left(F = \frac{e e'}{K r^2}\right)$ the quantity K is a constant for any one medium, and so it is called the "dielectric constant." Since in all conducting media a strain is impossible, K must for them be infinitely great. Its value for various dielectrics will be given later.

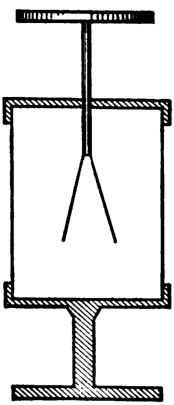


FIG. 146.

Electroscopes. The fact that two bodies which are similarly charged repel each other may be made use of in an instrument designed to detect charges. One form of instrument is called the "gold-leaf electroscope." It consists of a metal plate to which is attached a metal rod carrying at its lower end two long narrow strips of gold-foil. These are suspended, as shown, in a glass jar, so as to protect them from disturbances. As just explained, a metal body of any shape has the property of allowing a charge placed at one point of it to spread over the entire surface. So, if a charge of any kind is given the top metal plate, it will spread over the gold-leaves; and, as they are similarly charged, they will repel each other and stand diverged. If, now, another similar

charge is added, the leaves will diverge still farther; while, if the charge is of the opposite kind, it will neutralize some of the existing charge, and the leaves will come closer together.

As will be explained later, in speaking of induction (Art. 230), it is not necessary actually to add the charge to the metal plate in order to produce the motions of the leaves. If the charged body is simply brought near, the leaves will diverge more or will collapse, depending upon the nature of the charges.

Frictional Machines. The general method of producing a charge is to rub together two different dielectrics, e. g. glass and silk, and then to separate them. If the charges are removed from the two bodies, fresh charges may be produced by repeating the process, etc. This succession of steps, rubbing, separation, discharge, may be made automatic, as it is in the so-called "frictional machines." The charges are removed from the charged dielectrics by passing conductors over them, or by allowing conductors provided with sharp points to pass *near* them. As will be explained later (see Art. 230), this last method also produces the discharge of the charged dielectric and the appearance of an equal charge on the conductor.

221. Energy of Medium. As stated before, when two bodies are charged by contact and separation, the energy which is associated with the charges is in the surrounding medium; and this space is sometimes called the "electric field." The amount of energy is equal to the work done in separating the charges, and some idea may be formed of relative amounts from a consideration of the formula $F = \frac{e e'}{K r^2}$. It is evident that the greater the charges, so much the more work must be done to separate them. Further, K is different for different media, being greater for almost all substances than for air; so more work is required to separate the charges in air than in other dielectrics.

Consequently, there is more energy per cubic centimetre in air than in other dielectrics under similar conditions. The amount of the energy in one cubic centimetre must also be greater near the charges than farther away, because r is small there. Since there is no strain at all inside a conductor, there is no energy inside; and so, as far as electrostatic energy is concerned, a conductor acts like a medium, for which K is infinitely great, as was noted above.

Motions in Electric Fields. Since all motions which take place of themselves do so in such a way as to cause a decrease of potential energy, it must be possible to "explain" thus all motions in electric fields.

Two bodies, one charged positively, the other negatively, attract each other, because work is required to separate them; and so, by coming closer together, the potential energy becomes less. (Compare a body falling to the earth.) Two bodies charged the same way repel each other; because, by so moving apart, the strain in the field becomes less.

If a piece of dielectric for which K is greater than for air (e.g. a bit of glass, paper, paraffin) is placed near a charged body in air, the immediate effect is to weaken the strain and so to decrease the energy in that portion of space occupied by the dielectric which is inserted. But, if this same piece of dielectric is placed nearer the charged body, where the strain would naturally be greatest, the decrease in the energy is greater than it would be in a position farther away. Consequently, since by the piece of dielectric moving nearer to the charged body the potential energy is decreased, the motion will take place of itself. In other words, a charged conductor, either + or —, will attract such a piece of dielectric in air. Or, in general, such a piece of dielectric, if placed in air, will tend to move from a place where the force is weak to one where it is strong. Of course, if K for the dielectric was less than for air, the motion would be in the opposite direction. By experiments

of this nature, the value of K, compared with its value for air, may be found for various dielectrics.

Since K for a conductor may be regarded as infinite, a piece of a conducting substance will be attracted by a charged body much more than any piece of dielectric. (Of course, to produce motion actually, the piece of matter to be attracted must be suspended perfectly free to move, or else must be very light in weight.)

222. "Lines of Force." An "electric field" has been defined as the name given to the space surrounding charged bodies, in which electric forces may be perceived. The properties of the field may be best represented by drawing certain lines in it which mark the direction of the force at each point in the field. A minute positively charged body, placed at any point in the field, will tend to move in some direction. If the body is so small as to be considered a "particle" (see Art. 27), this direction is called the direction of the force at that point where it is placed. A "line of force" is such a line that at each of its points its direction is that of the force at that point. (A charged particle, if left free to move, would not continue to move along a line of force; owing to the inertia of the matter composing the particle, it would at any instant have a certain momentum, and the actual direction of its motion would be determined by the geometrical sum of this momentum and that produced by the electric force.) It should be particularly noticed that lines of force cannot be drawn in conductors, because in a conductor there is no force if the charges are not changing, as already explained. So lines of force can be drawn only in dielectrics. But, further, inside a hollow closed conductor no lines of force can be drawn; for a line of force must start from a positive charge, and, as explained before, there are no charges inside a closed conductor.

Thus, lines of force always start from positively charged bodies, and end on negatively charged ones. If the field of

force is due to a positively charged spherical conductor (the equal negative charge being removed to an infinite distance), the lines of force are straight lines drawn perpendicular to the surface in all directions.

Other illustrations of lines of force are given in the accompanying diagrams; they are taken, almost without change, from J. J. Thomson's "Elements of Electricity and Magnetism," and the lines are so drawn as to represent by their numbers the relative charges on the different bodies. The first represents the lines of force due to two equal and opposite charges.

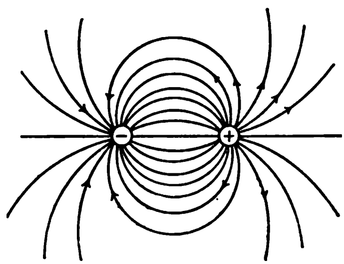


FIG. 147.

The second represents the lines of force due to two equal charges, eithor positive or negative. If they are positive, the lines are leaving the bodies; if they are negative, they are ending on them. (The other ends of the lines are on the equal opposite charges which are supposed to be off at an infinite distance.)

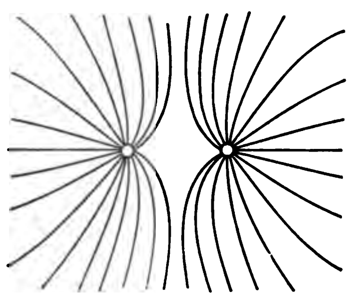


FIG. 148.

The third represents the lines of force due to a positive charge four times as great as the negative charge.

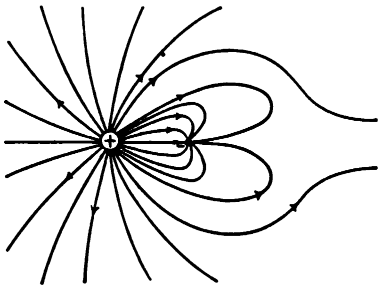


FIG. 149.

The fourth represents the lines of force due to a charge at A, four times as great as a similar charge at B.

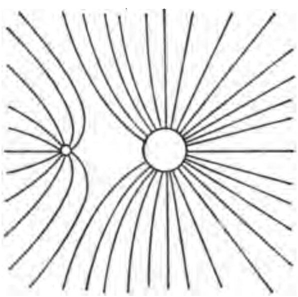


FIG. 150.

The fifth represents the lines of force due to a charge on a conductor formed of two spherical surfaces cutting at right angles.

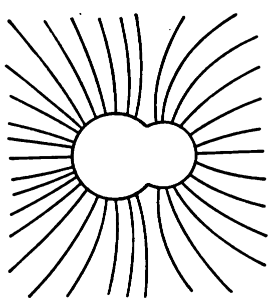


FIG. 151.

The sixth represents the lines of force due to two equal and opposite charges placed on two parallel conducting planes.

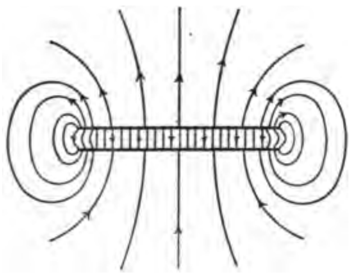


FIG. 152.

In this last it is evident that the strain is almost entirely in the space between the parallel plates; the charges have arranged themselves as they have, so as to make the potential energy as small as possible. In this space, too, the lines of force are parallel at points away from the edges of the plates; and such a field of force is called "uniform."

By studying these diagrams, it is seen that, where the lines of force are most numerous, there the electric force is strongest; and a great deal may be learned about the properties of any field by a consideration of the lines of force. One thing is evident: as two oppositely charged bodies approach each other, the lines of force become shorter; and so it is sometimes said that "lines of force tend to contract;" but this is, of course, merely a description.

CHAPTER II

ELECTRIC POTENTIAL AND INDUCTION

IN order to discuss more fully the properties of electric charges, it is necessary to introduce some mathematical ideas and definitions. As was seen in Chapter I., the most important laws of charged bodies depend upon relative amounts of potential energy; and the same statement is true of the charges themselves.

223. Electric Potential. The amount of external work necessary to move a *unit positive* charge from a point P in

FIG. 158.

the field to a point Q is called the "difference of potential" between Q and P. If this difference of potential is represented by E, the work done in making a positive charge, e, pass from P to Q is eE. As a result of this work, an equal amount of potential energy is given the surrounding dielectric. E may be a positive or a negative quantity. If it is positive, work is actually done by some external cause, and energy is added to the medium. If it is negative, the medium loses energy, and the change takes place of itself, the energy being spent in giving kinetic energy to the body which has the charge, or in doing some external work. Thus, if there is a positive charge in the

field at A, and if Q is nearer A than P is, the difference of potential E is positive, because work must be done by external forces to make a positive charge pass from P to Q. If there is a negative charge, though, at A, the difference of potential E is negative, because a positive charge would move of itself from P to Q.

Another definition of the difference of potential E is that it is the amount of work which a unit *positive* charge will do if it moves of itself from Q to P. Still another is that E is the amount of work required to carry a unit *negative* charge from Q to P.

These amounts of work, or changes in potential energy, are entirely independent of the path followed in moving between P and Q. (If it were not so, it would be possible to get "perpetual motion.")

224. Electric Forces. Sparks. If E is positive, external work is required if a unit positive charge is carried from P to Q; and so a force must have been overcome. The direction of the force at any point has been defined as the direction in which a *positive* charge would move as a result of the force, if placed at that point; and so the direction of the force in this case between P and Q is *from* Q towards P, since E is positive. If E is negative, the direction of the force is from P to Q, because a positive charge would move of itself in that direction. If $E = 0$, i. e. if there is no difference of potential, no work is required to pass from P to Q, or *vice versa*. This implies that there is no force opposing or helping the motion. If there is a force, there must be a difference of potential.

If the difference of potential between two points is great, it is equivalent to saying that the electric force is great between the two. Consequently, the strain in the dielectric of the field must be great. This condition of affairs may be secured by placing near each other two bodies, one charged positively, the other negatively. There is now a great difference of potential between two points, one on

each charged body; and there is a great strain in the intervening dielectric. If the strain is too great, there is a rupture of the dielectric, a spark passes, and the two charges combine. A spark acts like a perfect conductor connecting the two bodies; it consists of gases which are made luminous by the rupture of the medium and the motion of the charges.

A spark will, then, pass, in general, at points where the strain is greatest; that is, where the difference of potential is greatest or the electric field most intense. If the dielectric is weakened at any point, a spark will, of course, pass there more easily.

225. Potential at a Point. As will be immediately proved in the following article, there can be no difference of potential between any two points of a conductor on which the charges are at rest. (If there were, there would be a flow of the charges.) A special case of a conductor is the earth; and so there is no difference of potential between any two points of it, if the charges are at rest. Owing, however, to its immense size compared with the conductors which are ordinarily used, any variations of the charges on the earth due to effects produced at a given point on its surface are too minute to be observed. So it may be said that there is never any change in the electrostatic condition of the earth. Further, since the remote side of the earth (still at the same potential as all nearer points) is so far away, it may be said that there is no difference of potential between it and infinity, i. e. a point at an infinite distance from us.

Consequently, it is convenient to measure all differences of potential from the earth to various points, that is, to use the earth as the starting-point. And the numerical value of the electric "potential at a point" is defined to be the amount of work required to carry a *unit positive* charge from the earth to that point, or from infinity to that point. This is equivalent to calling the electric potential of the

earth and of infinity zero; that is, the electric potential of the earth is taken as a fixed point. (This is done arbitrarily, just as the temperature of melting ice is called 0° in the Centigrade system.)

226. Potential of a Conductor. The potential of all points of a conductor, inside and on the surface, must be the same, if the charges are at rest; for, if there was a difference of potential, there would be a force in the conductor, which is impossible, since the charges are at rest.

Further, the potential at any point inside a hollow closed conductor is the same as at a point of the conductor itself. For it has been proved (Art. 220) that there are no charges inside a hollow closed conductor, and that consequently there are no lines of force inside. And, if there was a difference of potential between a point in the hollow space and a point of the conductor, there would be a force between them; and so a line of force could be drawn. But this has just been shown to be impossible. (If there is a second charged body placed in the hollow space, but insulated from the conductor, this last statement does not apply.)

As will be shown later, the potentials of all points of a conductor are not in all cases the same, even if the charge is at rest. If the temperature of a charged conductor is not the same at all points, or if two conductors of different materials are in contact, the potential is different at different points; but these variations are exceedingly minute in comparison with the potentials of ordinary charged bodies.

227. Distribution of Charges on Conductors. As just stated, when a charge is placed at any point on a conductor, it distributes itself over the surface in such a way that the potential of the surface is the same at all points, and so that there is no force at any point inside the outer surface. The charge per square-centimetre of the surface, what may be called the "surface-density" of electricity, is by no means the same all over the surface. It is much

greater at sharp points and edges than on flat surfaces. This is made evident by the ease with which charges pass from points off to the air. In the neighborhood of points, where the surface density is great, the electric force must be large; and, if it is sufficient to strain the air near the points so that it breaks down, a minute spark is caused, and some of the charge passes from the sharp point of the conductor to the air. The air is now charged with the same kind of electricity as the conductor, and so the two repel each other. On a smooth sphere the surface-density would be the same at all points.

228. Equipotential Surfaces. There must be in any electric field many points where the potential is the same; in other words, such points that no work is required to carry a charge between them. The locus of all points whose potentials are the same is called an "equipotential surface." (Compare the fact that no work is done against gravitation in moving any body along a horizontal table. The table is then a gravitation equipotential surface.) It is possible to construct a series of equipotential surfaces, corresponding to different values of the potential, each value being a constant for one surface. Thus, the surface of a charged conductor is an equipotential surface, if the charge is at rest. Again, if the charge is on a spherical conductor, the equipotential surfaces are evidently, by symmetry, concentric spherical surfaces.

The direction of the electric force at any point of an equipotential surface must be perpendicular to it. If it was not, there would be a component of the force in the surface; and in carrying charges along the surface, work would be required in order to overcome this component. But this is impossible, because by definition there is no difference of potential in the surface. Consequently, lines of force cut the equipotential surfaces at right angles. Sections of the equipotential surfaces may thus be drawn for all the charges represented in Article 222, by simply

drawing closed curves perpendicular to the lines of force as given.

Further, the direction of the lines of force is always from high to low potential. Let, in the figure, V_1 and V_2 be partial sections of equipotential surfaces, and f be the line of force perpendicular to them. In carrying a positive charge from V_2 to V_1 along the line of force, work must be done, if the direction of the force is from V_1 to V_2. Hence $V_1 - V_2$ is positive; and so V_1 is greater than V_2.

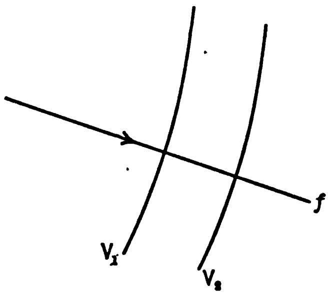


FIG. 154.

229. Induction. If there is a positive charge producing the field, the potential at a point A, which is near it, must be greater than that at a point B, farther away, because

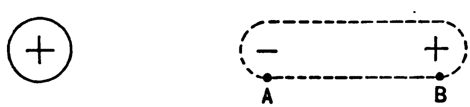


FIG. 155.

work would be required to carry a positive charge from B to A. If an uncharged conductor of any shape is now placed connecting A and B, there must be some change so as to make the potential of A and B the same. The potential at A must be lowered and that at B raised. This can be done only by the appearance of a negative charge at A and a positive one at B. As noted before, any neutral conductor may be considered as having on it equal amounts of positive and negative electricity; and so the amount near A is equal and opposite to the amount near B.

Similarly, if the field is produced by a negative charge, the potential at a point A, near the charge, is less than at a

point B, farther away, because work is required to carry a positive charge from A to B. And if an uncharged conductor is placed so as to connect A and B, the potential of A must be raised by the appearance of a positive charge there, and that at B lowered by an equal negative charge.

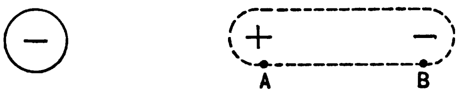


FIG. 156.

These charges which thus appear on uncharged conductors when they are placed in electric fields are called "induced" charges; and the entire phenomenon is called "electrostatic induction."

The accompanying drawing gives the field of force after an uncharged spherical conductor is placed in a uniform field.

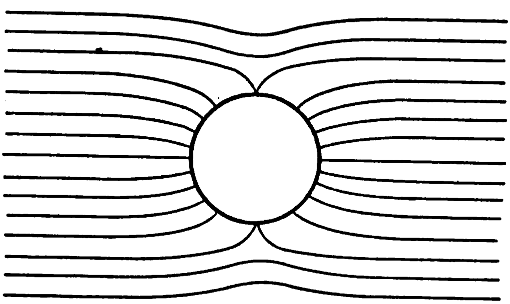


FIG. 157.

Where the lines of force end on the conductor, there is a negative charge; and, where they leave the surface, there is a positive charge. Where the lines of force are most numerous, the field is, of course, most intense.

A drawing is also given of the change produced in a uniform field of force in air by the insertion in it of a sphere of some dielectric like glass or paraffin, for which K is greater than for air.

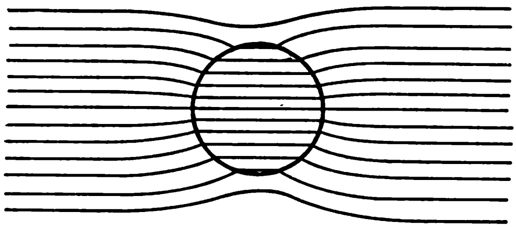


FIG. 158.

The field of force inside the sphere is also uniform (it would be so also for any ellipsoid placed with its axis parallel to the original field). The field of force is more intense just outside the sphere than at a long distance away; so any small pieces of uncharged conductors or dielectrics would be attracted up towards the sphere. (See Art. 221.)

It requires energy of course to produce these induced charges; and the work is done by the external force which brings the uncharged conductor into the field of force, or which brings the field of force near the uncharged body.

230. Illustrations of Induction. The fact that when a charged body is brought near a gold-leaf electroscope, changes in the position of the leaves are produced, is at once explained as due to the induced charges caused by the charge which is carried near.

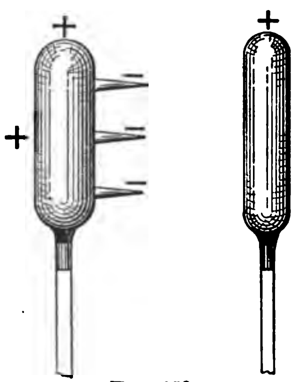


FIG. 159.

Again, the reason may be given why a conductor which has on its surface sharp points will discharge any charged body near the points, and will itself become similarly charged. The charged body induces charges in the conductor, — an opposite charge on the pointed side nearest it and a similar charge on the further side. As previously explained, the charges on the

points may easily pass off to the air; and then these charged particles in the air will be attracted by the originally charged body, since the charges are opposite. So the latter body becomes discharged, while a charge has remained on the pointed conductor.

231. "Charging by Induction." If a conductor is placed in a field of force, its potential will be that of the field at the point where it is placed; and this will be, in general, either positive or negative, that is, greater or less than that of the earth. So, if the conductor is now joined to the earth by some conductor, e. g. a wire or the human body, its potential must immediately become 0. If its potential was + before, then a negative charge must appear on it so as to lower the potential (the equal negative charge is on the earth). If its potential was − before, a positive charge will appear on it so as to raise its potential to 0. (This case is just like the one treated in the previous section, Article 229, if the point B is taken on the earth, and the conductor is supposed to consist of the original uncharged conductor, the wire and the earth.) If now the connection with the earth is broken, the conductor which was originally uncharged is charged either positively or negatively, depending upon its original potential. This process of producing charges on bodies is called "charging by induction."

232. Faraday's "Ice-Pail Experiment." A famous experiment in induction, and one of great theoretical importance, was performed by Faraday by means of an ice-pail and a small charged conductor. It may be repeated, using instead of the ice-pail a hollow conductor, which is more nearly closed. The nearly closed hollow conductor is insulated from the earth, and is joined by a wire to an electroscope, such as a gold-leaf one. The conductor is uncharged, and so there is no diverging of the gold-leaves. If a small charged conductor is now carefully lowered by means of a silk thread into the interior of the hollow

conductor, so as not to touch it at any point, charges are, of course, induced on the inside and outside of the hollow conductor. Consequently, the gold-leaves of the electroscope will diverge. Two interesting facts may be observed: (1) if the small charged conductor is moved at random inside the hollow conductor, but not touching it, there is

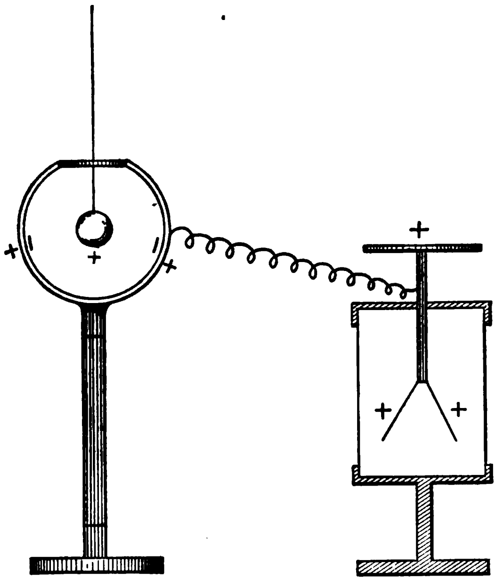


FIG. 160.

no change in the divergence of the gold-leaves; (2) if the small charged conductor is allowed to strike the inside of the hollow conductor, there is still no change in the leaves. This proves that the small charged body induces on the inside and outside of the practically closed conductor around it charges equal to itself, — one opposite, the other similar. The two induced charges must, of course, be equal to each other but of opposite kinds. The charge on the inside must be of an opposite kind to that on the small body which is inserted; and these two must be equal; because,

when the two conductors touch, there is no change on the outside (as proved by the absence of change of the gold-leaves), and so the charge on the small conductor must exactly neutralize that on the inside of the hollow conductor.

233. "Faraday Tubes." Faraday proposed a description of these phenomena of induction, which has the great advantage of presenting a clear picture of them to the mind. He thought of there being tubes constructed throughout an electric field, these tubes being built up of lines of force; that is, a closed curve is taken in the field of force, and lines of force are drawn through each point of the curved line, thus making a hollow tube. Each tube is made of such dimensions that, when it reaches a charged body, its end shall exactly include a unit charge. Of course, different tubes have different sizes; and the cross section of any one tube varies, except in a uniform field, where the tubes are of the same cross-section throughout.

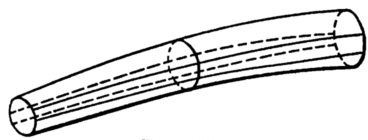


FIG. 161.

Thus, a body charged with 6 units has six tubes leaving it, if it is +; or ending on it, if it is negative. And the quantity of the charge on any body may be known if the number of tubes ending on it or starting from it is known. In particular, when an uncharged conductor entirely encloses a body which is charged with a quantity $+e$, e tubes leave the inner body and end on the inner side of the enclosing conductor. Hence there is a charge $-e$ on the inside of this conductor. But since this conductor was originally uncharged, as many tubes must leave it as end on it; and so $+e$ tubes must leave the outer surface, showing that there is a charge $+e$ there.

Again, where the surface density of a charge on a conductor is greatest, there must be the greatest number of Faraday tubes.

234. Shielding by Closed Conductors. As stated above, in Article 226, the potential of a closed conductor is the same at all points inside, so far as any charges on the outside of the conductor are concerned. So there is no electric force in the interior; for, if there was, there would be a difference of potential. Consequently, no matter what electric changes go on outside a closed conductor, there is no corresponding change inside; so that the interior is entirely shielded from external effects. (This is not true of the magnetic effect produced by electric currents.)

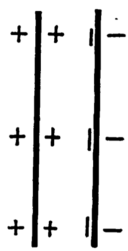
Again, if there are any charged bodies in the interior of a closed conductor, there will be charges induced on both the inner and outer surface of the conductor; and so there will be a field of force outside, which will change if the charges inside are changed. But, if the conductor is joined to the earth by a wire or other conductor, the field of force outside vanishes; because, now both the conductor and the earth are at the same potential, and so there are no lines of force in the field. (A line of force must pass from high to low potentials.) Consequently a hollow closed conductor joined to the earth completely shields the exterior space from any effects due to charges inside.

Induction-Machines. Various devices have been made by means of which to utilize the method of charging by induction so as to secure unlimited amounts of electricity. These are called "induction-machines." The oldest one is known as the "electrophorus," and was invented by Volta. Later forms are the Voss machine and the Wimshurst machine. Full descriptions of these are given in larger treatises on Experimental Physics. The principle made use of in them all is to turn mechanical energy into electrical by using mechanical power to produce induced charges.

235. Condensers. If a conductor is placed in air at some distance away from other conductors, and if it is given a succession of positive charges, its potential will rise, be-

cause more and more work will be required to carry a positive charge from the earth to it. (Similarly, if it was given a constantly increasing negative charge, its potential would fall.) A limit will finally be reached, though, when for a certain charge the difference of potential between the conductor and the earth (or some other conductor) is so great that a spark will pass; that is, the surrounding dielectric is so strained that it yields to the stress and "breaks down." If, however, the potential of the charged conductor could be made nearer that of the earth in any way, the strain would be lessened; and so the conductor could contain a greater charge before the medium would break.

This can actually be done. Let A be a metal plate charged positively, and let its potential be V; the problem is to see how this potential can be diminished without decreasing the charge on the plate. Bring near A an insulated uncharged conductor B, such as another metal plate; charges are induced on B, negative on the side towards A, positive on the further side. If a unit positive charge is now brought up to A from the earth, less work is required than was before B was in place, because now most of the strain in the medium is between A and B (or, in terms of another description, because the negative charge on B, being nearer A than the positive charge is, helps in doing the work by its attraction on the positive charge which is being brought up from the earth). Consequently the potential of A has been lowered. It may be lowered still further if the conductor B is joined to the earth by a wire or other conductor, because now still more of the strain is drawn into the space between A and B (or because the positive charge on B now goes to the earth); and so less work would be required to bring up a unit positive charge from the



A B

FIG. 162

earth. The potential of the conductor B is now 0, the same as that of the earth; and that of the conductor A equals the amount of work required to carry a unit positive charge from B to A. If a plate of any dielectric for which K is greater than air is now inserted between A and B (e. g. a plate of glass or of paraffin), replacing the air, less work would be required to carry a unit positive charge from B to A, because in glass or paraffin electric forces are less than in air, there is less strain than in air. So the potential of A is again lowered. To sum up, then, the potential of A may be lowered by bringing near it another conductor connected to the earth and separated from A by a dielectric for which K is as great as convenient. So that now the conductor A can receive a much greater positive charge than it could when by itself, before its potential is raised to any definite value.

In an exactly similar manner it may be proved that, if A is charged negatively and so has a negative potential, its potential may be *raised,* i. e. brought nearer that of the earth, by bringing near it a conductor B joined to the earth and separated from A by a dielectric, for which K is very large. So the conductor A must now receive a much greater negative charge than it had before, if its potential is to be as low.

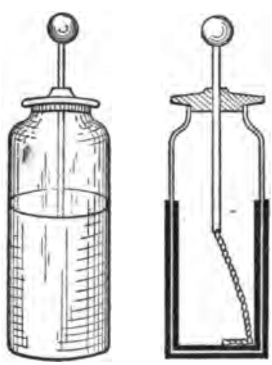


FIG. 163.

Such an arrangement, two conductors separated by a dielectric, is called a "condenser." The commonest type of condenser is one where the conductors are parallel plates; these may be flat pieces of tin-foil separated by flat pieces of glass or paraffined paper; or they may be curved, as in the Leyden Jar, which is a glass bottle, coated inside and out with tin-foil, as shown. The tin-foil does not

extend up to the edge of the bottle, and the interior foil is connected to a metal ball outside by means of a metal chain.

If one of the conductors, B, is joined to the earth, its charge will be opposite to that on A; and, if the two conductors are near each other, like two parallel plates close together, or if one conductor encloses the other like two concentric spheres or two coaxal cylinders, the two charges are almost exactly equal. (In the case of two concentric spheres, they are equal.) For, practically, all the Faraday tubes leaving one conductor end on the other. If the connection with the earth is now broken, and the entire apparatus moved elsewhere, the potentials of the two conductors may change, but their difference will remain the same, because the same amount of work is required after the change in position as before, to carry a unit positive charge from one conductor to the other. Consequently a charged condenser may be considered as made up of two conductors at different potentials, V_1 and V_2, carrying charges $+ e$ and $-e$, and separated by a dielectric whose constant is K.

236. Discharge of Condensers. The medium between the two conductors of a charged condenser is of course strained, and the discharge of the condenser consists in the release of this strain. The simplest method is to join the two conductors by a wire; for then their potentials become the same, the two equal and opposite charges neutralize each other, and the strain disappears. Analogy from mechanics would lead us to expect two kinds of discharge, a steady one and an oscillatory one. If a wire strained by twisting is allowed to untwist and so lose its strain, it can do so in two ways, depending upon external conditions; if the wire is held rather firmly by the fingers, but is allowed to slip through them, the twist disappears slowly and continuously; if the wire, though, has considerable inertia and is free to twist, it will untwist, twist in the opposite direction,

untwist, etc., making a series of oscillations or reversals, but finally coming to rest as they die out. That is, if there is great frictional resistance, the discharge of the strain is continuous; if there is great inertia and no friction, the discharge is oscillatory.

It is exactly the same in the discharge of a condenser; the medium is strained and held so by definite forces; when these forces are removed, the discharge takes place. If there is a great frictional resistance to the dying out of the strain, it will discharge continuously, the positive charge will grow less, and so will the negative; if there is a small frictional resistance, the strain will discharge itself, then become reversed, discharge itself, return to its previous type, etc. When the strain is reversed, the conductor which was positively charged becomes negative, and the one which was negative becomes positive; but, as the strain is repeatedly reversed, the corresponding charges become less and less, and so in the end the strain is discharged. The first of these two kinds of discharge may be obtained by joining the two conductors of the charged condenser by a bad conductor; the second, by joining them by a good conductor. The discharge is in any case extremely rapid, and is studied with difficulty.

In some condensers it is observed that, after being apparently discharged by having their two conductors joined, they become charged again on standing some time; and, if this secondary charge is discharged, there may be another charge, etc.; but each succeeding discharge is much less than the preceding one. It has been proved that these secondary discharges occur only in such condensers as have for their dielectric some heterogeneous material such as glass. A condenser which has a gas, a liquid, or a homogeneous solid like a pure crystal for its dielectric has no secondary discharges. When a condenser with a non-homogeneous dielectric is discharged, the strain does not vanish, but some parts are strained in opposite directions

to other parts, thus producing apparent absence of strain. But, if now the medium yields to the strain in some part, the opposing strains are no longer balanced, and so there is a secondary charge.

237. Capacity. If a given conductor is placed by itself in some dielectric, at a great distance from other conductors, and is given a charge of any kind, it will have a definite potential. Let the charge be e, and the potential be V. It is found by experiment that one is proportional to the other; that is, that, if there is twice the charge, the potential is twice as great, etc. This may be expressed mathematically by writing

$$e = C\,V. \quad \quad (1)$$

C is a constant for the particular conductor, and the particular surrounding dielectric. C does not depend upon the material of the conductor, but only on its shape and size. It is called the "capacity," and is evidently the numerical value of that charge which is required to produce a potential 1 on the particular conductor in the particular dielectric. It may be proved that the capacity of a spherical conductor of radius R in a dielectric whose constant is K is $C = K\,R$.

Similarly, if the equal and opposite charges on a condenser are varied, the corresponding differences of potential vary in such a way that, if e is the charge on one conductor ($-e$ being on the other) and $V_1 - V_2$ the corresponding difference of potential,

$$e = C\,(V_1 - V_2), \quad \quad (2)$$

where C is a constant for the particular condenser, and is called its capacity. C depends obviously on the shape and size of the two conductors and upon the dielectric between them.

For two concentric spheres of radii R_1 and R_2, separated by a medium whose constant is K, $C = \dfrac{K\,R_1\,R_2}{R_1 - R_2}$.

For two parallel plates of area A, at a distance d apart, and separated by a medium whose constant is K, $C = \frac{K A}{4 \pi d}$ ($\pi = 3.1416$, the ratio of the circumference of a circle to its diameter).

If any quantity of electricity is given to two condensers of identically the same dimensions, but having different

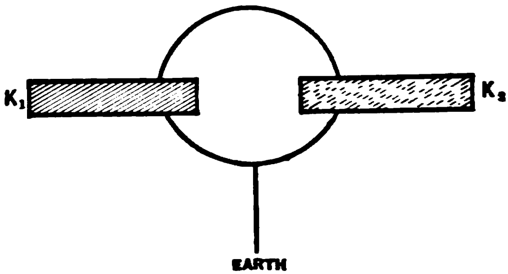


FIG. 164.

dielectrics, the charges will be divided in the ratio of the values of K. For let two such parallel-plate condensers be joined, so that one plate of each is connected by a wire to one of the other; and join one of these pairs by a wire to the earth. Give to the other pair of plates any quantity, e. This will distribute itself so that one condenser has e_1; the other, e_2 where $e = e_1 + e_2$.

But
$$e_1 = C_1 (V_1 - V_2),$$
$$e_2 = C_2 (V_1 - V_2),$$

since the differences in potential are the same. Hence $e_1 / e_2 = C_1 / C_2 = K_1 / K_2$, since the two condensers are identical, with the exception of the dielectrics. Thus, by comparing the values e_1 and e_2, the ratio of K_1 to K_2 may be determined. Some values of K are given in the following table:—

TABLE XII

DIELECTRIC CONSTANTS (ELECTROSTATIC SYSTEM).

Glass (about) .	6	Turpentine . .	2.4
Mica	8	Petroleum . .	2.1
Paraffin . . .	2	Hydrogen . .	0.9998
Rubber . . .	2.5	Illuminating Gas	1.0004
Water	76	Carbon Dioxide	1.0008
Alcohol . . .	26	(Vacuum) . .	0.9995

238. Energy. The energy which is associated with electric charges is, as has been said many times, in the dielectric forming the electric field; and it is not difficult to calculate the amount of the energy per cubic centimetre at any point of the field, expressed in terms of the properties of the field at that point. It is much easier, however, to calculate the entire energy of the field in terms of the charges and their potentials.

If there is a conductor whose charge is e and whose potential is V, the work required to produce this charge is $\frac{1}{2}\, e\, V$. The charging may be considered as having been done gradually, so that the potential has slowly risen from 0 to V. If the potential at any instant is V', the work required to bring up from the earth a unit positive charge is V', if the potential does not change; and so the work necessary to bring up a charge $+e$ would be $e\, V'$, if the potential did not change. But while e is being brought up in small equal amounts, the potential rises regularly from 0 to V, and so the average potential is $\frac{1}{2}\, V$. The entire work done, then, in bringing up the charge $+e$ to the average potential $\frac{1}{2}\, V$ is the product of the two, $\frac{1}{2}\, e\, V$. Since $e = CV$, this may be expressed

$$\text{energy} = \tfrac{1}{2}\, e\, V = \tfrac{1}{2}\, CV^2 = \tfrac{1}{2}\frac{e^2}{C} \quad . \quad . \quad . \quad (3)$$

(The negative charge $-e$ is on the earth at the potential 0; so this may be said to be the work required to separate the

charges $+e$ and $-e$ until their difference of potential is V.)

The charges on a condenser are $+e$ at a potential V_1 and $-e$ at a potential V_2. Hence the work required to produce these charges, that is, the entire energy of the field, is $\frac{1}{2} e V_1 - \frac{1}{2} e V_2 = \frac{1}{2} e (V_1 - V_2.)$ But $e = C (V_1 - V_2)$ and so this may be expressed,

$$\text{energy} = \tfrac{1}{2} e (V_1 - V_2) = \tfrac{1}{2} C (V_1 - V_2)^2 = \tfrac{1}{2} \frac{e^2}{C} \quad . \quad . \quad (4)$$

239. Electrostatic Measurements. There are three quantities which must be measured in order to express the electric properties of a charged conductor, — quantity of charge, potential, capacity.

Capacity depends upon the size and shape of conductor and the constant of the dielectric; and in certain simple cases its value may be calculated, as indicated in Article 237.

Potential, or difference of potential, may be measured by certain instruments known as "electrometers." One form of instrument depends upon the fact that the two plates of a parallel-plate condenser tend to approach each other because one is charged positively and the other negatively. It may be proved that, if

$V_1 - V_2 =$ difference of potential of the two plates,
$d =$ the distance apart of the two plates,
$A =$ area of a movable portion of one plate near its centre,

the mechanical force which must be applied to this movable disc to keep it from moving towards the other plate is

$$F = \frac{(V_1 - V_2)^2 A K}{8 \pi d^2}.$$

Hence

$$V_1 - V_2 = d \sqrt{\frac{8 \pi F}{KA}} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

If air is the dielectric, $K = 1$ on the electrostatic system; and it is possible to measure F, d and A; so $V_1 - V_2$ may be calculated.

A disc is taken near the centre of one plate, so as to have a region where the field is uniform. If the force on the entire plate was measured, it would be necessary to make a correction for the lack of uniformity at the edges.

Knowing capacity and difference of potential, quantity may be calculated. There is a method which will be described later on by means of which quantity can be measured, but not in terms of the electrostatic system.

CHAPTER III

ELECTRIC CONDUCTION

In the previous chapters, the general properties and laws of electric charges at rest have been considered. But it is possible for the charges to move, i. e. for the electric strain to change; and in this and the following chapters will be studied the different methods of producing motion of the charges, and the laws of the various phenomena associated with this motion.

240. Electric Current. If a series of small charged bodies move in a given line, or if a charge passes along a conductor, this phenomenon is called an electric "current." If a charge $+e$ goes by a fixed point in one direction in *one second*, and if in the same time a charge $-e'$ goes by in the opposite direction, the sum, $e + e'$, is called the "intensity" of the current, and its direction is said to be that in which the positive charge moves. If the current is steady, the quantity of electricity carried by in t seconds is the product of t and the intensity of the current.

The simplest type of a current in a conductor is that afforded when two plates of a charged condenser are joined by a conductor, e. g. a fine wire, as shown. The charges on the two plates become less and less, as is manifest by the vanishing of the strain in the medium between the two plates. While this process is going on, the wire joining the two plates grows warm, showing that its molecules are receiving energy; other effects are also produced, which will be discussed later; and the existence of a current can always be proved by any one of these effects. The

energy of the dielectric has thus passed into the wire; and there is evidence for believing that it passes into the wire perpendicularly, as shown by the dotted lines.

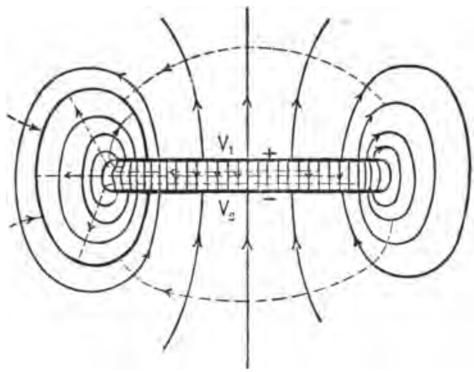


FIG. 165.

There is said to be an electric current in the wire joining the plates, as long as the energy continues to come into the wire from the dielectric. Unless the charges on the plates are renewed, this process will soon stop. But if, as fast as the strain dies down, it is restored, the process will continue unchanged; and there is said to be a "steady" current in the wire. This may be expressed mathematically by saying that, if the difference of potential, $V_1 - V_2$, can be maintained constant, there will be a steady current, and its direction in the connecting wire will be from the point of high potential to that of low, i. e. from V_1 to V_2. (Notice the analogy of the flow of heat energy from high temperature to low, and of the flow of volume energy from high pressure to low. In this case electric energy passes from high to low potential.) Wherever there is a difference of potential, there is an electric force from high to low; that is, a positive charge, if free to move, would pass from high to low potential, while a negative charge would pass in the opposite direction. So

in this case of the condenser, since a conductor joins the two plates, and since they may be considered as having equal amounts of positive and negative charges, the former will move in one direction in the wire, the latter in the opposite. While the charges are moving there is no reason for believing that they are on the surface of the conductor; and, in fact, it is known that, except in the case of rapidly alternating currents, the current is inside the conductor as well as on the surface.

The one necessary condition, therefore, for the production of a current in a conductor is some means of keeping its two ends at a difference of potential: if this difference is constant, the current is steady; otherwise, not. A name has been given to the difference of potential between any two points of a conductor through which a current is flowing, viz., the "electro-motive force" (E. M. F.) between those points.

There are many ways by which a continuous E. M. F. may be produced and maintained. The simplest method is to use an electric machine, either a friction or an induction one, and to join its two "poles" to the conductor in which a current is desired.

Another method is to make a circuit, part of one conductor, part of another; and to keep the two junctions at different temperatures. It is observed that under these conditions there is, except in special cases, a current, which is called a "thermo-electric current."

Again, it is observed that, if two different solid conductors dip into a liquid conductor, there is an E. M. F. between the ends of the solids, which rise above the liquid. So, if they are joined by a conductor, e. g. a wire, there will be a current in it.

Still another method depends upon certain magnetic properties, and will be discussed later.

In all these cases energy passes from the dielectric into the conductor, in which the current is "flowing;" and this

energy must be maintained by a supply of energy from some source. In the case of both the electric and magnetic machines, the work is done by some external body turning the machine. In thermo-electric currents the energy comes from the bodies which maintain the differences in temperature at the junctions. Where the two solid conductors dip in a liquid one, the energy comes from certain chemical reactions.

241. Thermo-Electric Currents. The conditions for a thermo-electric current are: a circuit made of two different conductors 1 and 2, having junctions at A and B; and the maintenance of different temperatures at these junctions. The explanation of the cause of the current is not difficult.

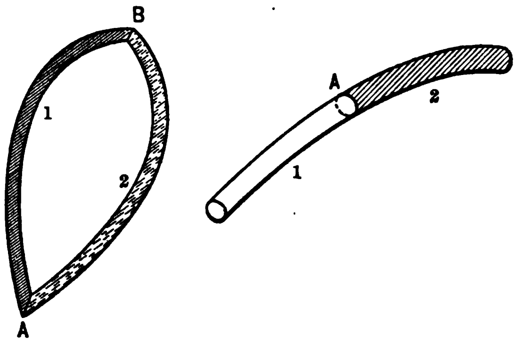


FIG. 166.

If any two conductors, 1 and 2, are placed in contact at A, they do not come to the same potential exactly (see Art. 226). Call the potential of conductor 1, V_1; and that of 2, V_2; and let $V_1 > V_2$. The natural tendency of all conductors in contact is to come to the same potential; but, since V_1 is higher than V_2, there must be some electric force acting from 2 towards 1, across A, so as to oppose the tendency of V_1 to become as small as V_2. A mechanical analogy may make the matter clearer. If at the bottom of a U-tube half filled with water there is a paddle-wheel, so

arranged as to drive the water in one direction or the other; and, if around the axle of this wheel a rope is coiled, at whose free end hangs a heavy weight, — the wheel will turn, thus forcing the water up in one arm of the tube to a greater height than it is in the other. The wheel will come to rest when the back-pressure due to the difference in pressure on the two sides of the wheel exactly equals the pressure which the wheel exerts owing to the heavy weight. The pressure at the side 1, p_1, is greater than that at the side 2, p_2, i. e. $p_1 > p_2$. These two pressures tend to become equal; but are kept from so doing by a pressure at A, produced by the falling weight, which acts from 2 towards 1 through A.

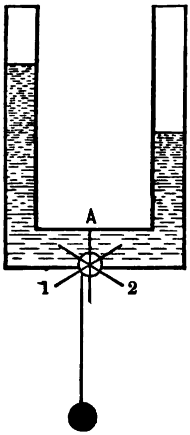


FIG. 167.

In the case of the two conductors 1 and 2, this force acting from 2 towards 1, i. e. tending to move a positive charge in that direction and a negative one in the other, must be produced by the differences in the "specific attractions" of the two kinds of matter for electricity. (See Art. 217.) Its existence may be easily proved. If a current is passed through the junction A from 1 to 2 by means of some electric machine (say), i. e. if a positive charge is forced from 1 to 2, and a negative one from 2 to 1, resistance will be experienced at the junction A if there is a force there acting from 2 to 1; work will have to be done there against "molecular forces;" energy will be given the molecules there; and so a heat effect which corresponds to addition of heat energy should be observed at the junction. The difference of potential is $V_1 - V_2$, or E. So this is the work which would be done if a *unit* positive charge was carried across from 1 to 2, or if a unit negative charge was carried from 2 to 1. If, then, a current whose intensity is

i, is carried across from 1 to 2 for t seconds, i. e. if a quantity $i\,t$ passes, the work done at the junction is $E\,i\,t$; and this work is spent in adding heat energy to the junction, since $V_1 > V_2$. Consequently, its temperature will rise, and this fact can be observed.

If, on the other hand, a current is forced through A from 2 towards 1, the electric force at the junction helps the current on, it itself does work; and so the molecules there lose energy, and the temperature of the junction falls. The loss in energy, if a current whose intensity is i passes for t seconds, is $E\,i\,t$; the same as the energy gained when the current is in the opposite direction.

Consequently, if on passing a current through any junction from 1 to 2 there is a rise in temperature; and if, on reversing the current, there is a fall in temperature, it is proved that there is a different of potential, i. e. an E. M. F. acting from 2 towards 1, so that V_1 is naturally greater than V_2. If the heat effect, measured in ergs, is H; and if the intensity of the current is i and it flows t seconds in producing H, the E. M. F., E, at the junction is given by the formula

$$H = E\,i\,t, \qquad (1)$$

i. e.

$$E = H / i\,t \qquad (1\,a)$$

This E. M. F. which exists at the junctions of any two different bodies (conductors or dielectrics), and which becomes evident whenever two bodies are brought in contact, is called the "Peltier E. M. F." (Its effect is often masked by the chemical action of surrounding media, e. g. air, upon the two substances.) It may be measured as just shown; and exact measurements prove that it changes as the temperature does, but at different rates for different pairs of substances.

It may also be proved by direct experiment that in a conductor, different points of which are at different temperatures, there are differences of potential. These are

called "Thomson E. M. F.'s"; and in some metals their directions are from points of high to points of low temperatures, and in others are in the opposite direction.

As a consequence of these two kinds of E. M. F.'s in a circuit of two conductors, whose junctions are at different temperatures, there are currents produced, unless the various forces balance one another. If the temperature of one junction is kept fixed, and that of the other varied, it is always possible to find such a temperature for it that there is no current. This second temperature is called the temperature of "inversion," corresponding to the fixed temperature of the other junction; for if the varying temperature is now changed so as to be farther away from the fixed temperature, the current is reversed. The energy which is required to maintain a thermo-electric current is furnished, as was said before, by the bodies which keep the temperatures of the junctions fixed.

By having a series of thermo-electric junctions, that is, by joining in series several "couples" of two conductors, and by keeping the alternate junctions at different temperatures, it is possible to produce a considerable E. M. F. The arrangement is shown in the figure; and if the junctions on the side A are at even a slightly different temperature from those on the side B, there will be a noticeable E. M. F., which can produce a current. Such an instrument is called a "thermopile," and it is extremely useful in detecting differences in temperatures.

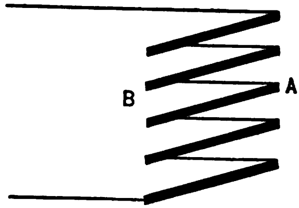


FIG. 168.

242. Primary Cells. If any solid dips in any liquid, there is, of course, a difference of potential between them; and this is found to be especially large if both the solid and liquid are conductors. The value of the difference of potential depends upon the nature of the two substances;

and, if two solid conductors dip in the same liquid conductor, they will as a rule be at different potentials. For the liquid conductor has the same potential throughout, and the potentials of the two solid conductors differ from that of the liquid by different amounts. So, if the solid conductors are now joined by a wire, there will be a current produced in it by the E. M. F. at its two ends, because this E. M. F. is maintained by the differences between the two solid conductors and the liquid. Such an apparatus is called a "primary cell."

A special case of this process is when rods of zinc and copper dip in acidulated water (e. g. a dilute solution of sulphuric acid). Both metals are found to have potentials lower than that of the liquid; but that of the copper is higher than that of the zinc, and so the E. M. F. produces a current from copper to zinc, to acidulated water, to copper, etc. It is observed, also, by experiment that, as the current flows, the zinc rod dissolves, or wastes away, and that at the copper rod hydrogen gas forms and bubbles off at the surface of the liquid.

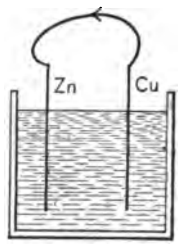


FIG. 169.

Before the zinc and copper rods are joined by the outside wire, the relative differences in potential may be represented by the diagram, in which vertical distances mean differences in potential. Thus, the zinc rod (Zn) has a constant potential, which is shown by a horizontal line; in passing from the zinc to the acidulated water there is a marked increase in potential, E_1; the liquid has a constant potential throughout, being a conductor; then in passing from it to the

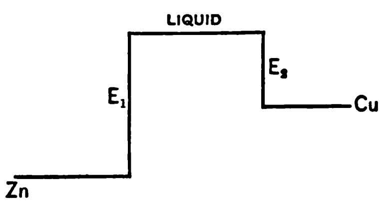


FIG. 170.

copper (Cu) there is a fall of potential E_2, but the copper is at a higher potential than the zinc, since $E_1 > E_2$.

If the copper and zinc rods are joined by a copper wire, a current is thus produced; and, since in any conductor the current flows from high to low potential, there must be a

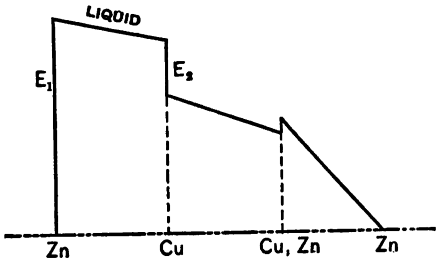


FIG. 171.

fall of potential in every conductor through which the current is passing, as is shown in the diagram. Starting from the side of the zinc rod next the liquid, there is a rise in potential to the liquid, a fall through the liquid, a drop in passing from the liquid to the copper, a fall through the copper rod and wire, a *slight* difference at the junction of the copper and the zinc, a fall through the zinc rod at a different rate from the fall through the copper, etc. These facts may be proved by direct experiment.

If the difference of potential between the zinc rod and the liquid is E_1, the work required to make a current whose intensity is i pass for t seconds from the zinc into the liquid is $E_1 i t$. Similarly, if the difference of potential between the liquid and the copper is E_2, the energy set free, as the current passes from high to low potential, is $E_2 i t$; that is, this amount of electric energy is transformed into some other kind of energy. In a similar manner, energy is set free throughout each conductor in the circuit, as the potential falls in the direction of the current. (The slight Peltier E. M. F. at the junction of the copper and zinc also requires or furnishes energy; but this amount is very

minute.) The energy set free by the current as it passes is spent in doing work at the junctions or in the conductors. As the current passes from the liquid to the copper, chemical actions take place, viz. hydrogen gas is formed, and the energy $E_2 i t$ is spent in doing this "chemical" work. Through the conductors there is a fall of potential, $E_1 - E_2$ (neglecting the Peltier E. M. F. at the Zn — Cu junction); and the energy $(E_1 - E_2)\, i t$ is spent in giving heat-energy to the molecules of the conductors. Consequently, in general, their temperature is raised. Therefore, the electric energy $E_1\, i t$ must be furnished from some source of energy in between the zinc and the acidulated water; $E_2\, i t$ is spent in chemical work; $(E_1 - E_2)\, i t$, in heat effects. At the zinc plate, as stated before, there is solution of the zinc in the acidulated water; and, as is well known, when zinc dissolves in acid, energy is liberated, so that the energy $E_1\, i t$, which is required for the current, is furnished by the molecules which dissolve from the zinc into the acid. (Some heat energy may be furnished from neighboring bodies, or may be given up to them; but these amounts are here neglected.)

If 1 gram of zinc is dissolved in a vessel containing acidulated water, hydrogen is evolved, and the temperature is raised. By measuring the rise in temperature the masses of the substances and their specific heats, it is possible to calculate how much heat energy has been produced when the 1 gram of zinc dissolves and the corresponding amount of hydrogen is produced. Let this amount of energy be h. Then if, as the current of intensity i passes for t seconds in the cell which has been described above, m grams of zinc dissolve and a corresponding amount of hydrogen is produced, the energy $m h$ is furnished as a result of the two chemical actions at the zinc and copper poles. Consequently,

$$m h = E_1 i t - E_2 i t = (E_1 - E_2)\, i t \quad . \quad . \quad (2)$$

So that, if m, h, i, t are known, $E_1 - E_2$ may be calculated.

A mechanical analogy may make the action of a primary cell clearer. If in a U-tube, partly filled with water, a pump or paddle-wheel is inserted at the bottom and worked by some external power, the water will be raised higher in one arm than the other; and, if the two arms are connected above by a pipe, the water may be forced over from one arm into the other. The process will continue indefinitely, as long as energy is furnished the pump so as to make it work. The pump forces the water from a low pressure on one side to a high one on the other, just as the chemical action at the zinc pole raises the electricity from a low potential to a high one.

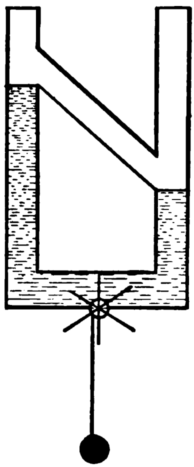


FIG. 172.

Such a primary cell as the one described is sometimes called a "single-fluid" cell, and the current which it furnishes is by no means constant. The hydrogen which is formed at the copper pole does not all bubble off; but some remains there, changing the character of the E. M. F.; some zinc is also slowly deposited on the copper. Again, as the positive charges are carried through the liquid, they are not all handed on to the copper pole, but some remain near it, thus opposing the approach of other positive charges. This last fact may be expressed by saying that there is an E. M. F. in the opposite direction to the one producing the flow through the liquid. The entire phenomenon is called "polarization," and the "back-E. M. F." is called the E. M. F. of polarization. To obviate these difficulties certain cells have been devised which contain two liquids separated by some porous partition, thus keeping the liquids apart, but permitting the current to pass. In these cells polarization may be largely avoided, and so a fairly constant current may be produced. If the current is too great,

the E. M. F. of any cell will change. The best-known cell of a two-fluid type is the Daniels. In this the two fluids are solutions of copper sulphate and sulphuric acid in water; and the two metals are zinc, dipping in the acid, and copper, dipping in the copper sulphate. Another cell, which is more complicated, but which maintains indefinitely a constant E. M. F., provided that only a minute current passes through it, is a Clark cell, which is now used in laboratories to give a standard E. M. F.

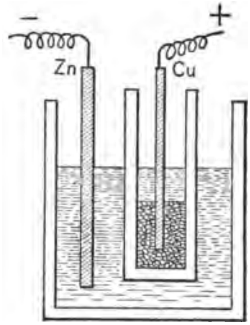


FIG. 173.

When a piece of pure metal is immersed in an acid, it does not continue to dissolve; on the contrary, the process ceases so quickly that it is detected with difficulty. But, if there is any metallic impurity in the metal, the process continues until the impurity is dislodged. The explanation is very simple. The metallic impurity, the metal itself, and the acid form a primary cell, like the copper-zinc-acid cell; and, as the current passes from the metal into the acid, the metal is dissolved, just as the zinc was. Ordinary commercial zinc rods are quite impure; so that, when they are dipped in dilute acids, some zinc dissolves without producing any current outside. This is called "local action;" and it may be largely prevented by amalgamating the surface of the zinc with mercury, so as to render the surface approximately uniform.

243. Nature of Conductors. All metals, in the solid and liquid conditions, are conductors; but only a limited number of other liquids are; and gases conduct only under definite conditions. There are certain complex conductors such that, when a current is passed through them, some of their constituents are liberated at the place where the current enters, and other constituents at the place where the

current leaves. Thus, when a current is passed through a solution of sulphuric acid in water, hydrogen and oxygen are liberated at the places where the current leaves and enters, respectively. Again, if a current is passed through a solution of silver nitrate in water, silver and oxygen are liberated. Such conductors are called "electrolytes;" and the process of the conduction through them is called "electrolysis."

244. Electrolysis. Some solids are electrolytes, such as silver iodide. With the exception of fused metals, all liquid conductors are electrolytes; and, when a gas is made conducting, it becomes an electrolyte. The metal conductors, sometimes called "electrodes" or "poles," by which the current enters and leaves a liquid or gaseous conductor, are called "anode" and "kathode," respectively.

The laws of conduction through electrolytes have been investigated by Faraday, and may be thus expressed:

1. The mass of the substance liberated in a given time at either kathode or anode is directly proportional to the quantity of electricity which has passed, i. e. to the product of the intensity of the current and the number of seconds, $i\,t$.

2. If the same current is passed through several electrolytes in series, i. e. if several substances are being liberated at the same time, the masses of the substances liberated are directly proportional to their "chemical equivalents." By definition, the chemical equivalent of any element is its "atomic weight" divided by its "valence." And by "valence" is meant the number of hydrogen atoms which combine with one atom of the element to form a stable compound; or, what is the same thing, *twice* the number of oxygen atoms which combine with one atom of the element to form a compound. In a similar way, the valence of any group of atoms or "radical" may be defined as the number of hydrogen atoms or twice the number of oxygen atoms which combine with it to form a molecule.

Thus, the valence of SO_4 is 2, because there is a molecule, H_2SO_4, sulphuric acid.

The following table gives the "atomic weight," valency, and chemical equivalent of several elements: —

TABLE XIII

	Atomic Weight.	Valency.	Chemical Equivalent.
Chlorine	35.37	1	35.37
Copper (Cupric)	63.18	2	31.59
Hydrogen	1.00	1	1.00
Iron (Ferric)	55.88	3	18.63
Lead	206.39	2	103.20
Oxygen	15.96	2	7.98
Potassium	39.03	1	39.03
Silver	107.66	1	107.66
Sodium	23.00	1	23.00
Zinc	64.88	2	32.44

Faraday's first law furnishes at once a method by means of which the intensities of two currents may be compared. If the mass m_1 is liberated by a current whose intensity is i_1 flowing for t_1 seconds, and if m_2 is liberated in the *same* electrolyte when a current of intensity i_2 flows for t_2 seconds,

$$m_1 : m_2 = i_1 t_1 : i_2 t_2,$$

or

$$i_1 : i_2 = m_1 t_2 : m_2 t_1 \quad (3)$$

And m_1, m_2, t_1, and t_2 can be easily measured. This principle is made use of in "gas voltameters," where gases are liberated; and in "copper and silver voltameters," in which copper and silver, respectively, are deposited at the kathodes. An illustration of a common type of gas voltameter is given on the following page.

245. From a study of electrolysis and its laws as dis-

covered by Faraday, a great deal may be learned as to the nature of conduction. The mechanism of the conduction is the passage of a series of positive charges in one direction and of negative ones in the other, these charges being carried on portions of matter; because there is no steady current unless portions of matter are liberated at the poles, and the amount of the current is directly proportional to the amounts of matter set free. The name given these portions which carry the charges is "ions." By examining at which pole, anode or kathode, a given substance is liberated, the kind of electric charge on the ions may be determined. If a substance is liberated at the kathode, i. e. where the current leaves, its ions must have been positively charged; and, if liberated at the anode, negatively charged.

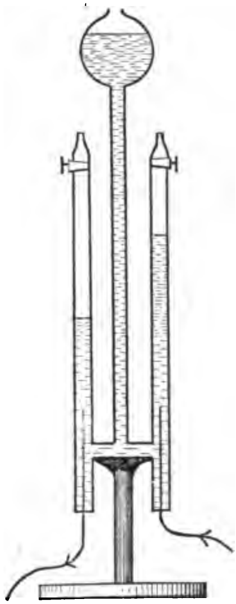


FIG. 174.

Experiments seem to prove that in every case of electrolysis the ions are dissociated portions of molecules. Thus, no gas conducts until some of the molecules are dissociated, — a state of affairs which can be proved to exist by other methods. Again, with the exception of fused substances, no pure liquid conducts, some salt or acid must be added to it; and there is a great deal of independent evidence leading to the belief that no such liquid is a conductor unless the salt or acid in solution in it is, at least partially, dissociated.

Faraday's second law fixes the relative charges carried by an ion of any form of matter. First, the charge carried on any one ion of the same substance is always the same, no matter in what electrolyte the ion exists. This is

proved by the following fact: if the same current is passed through two electrolytes in series, which liberate at their two similar poles, e. g. kathodes, the same substance (such as solutions of sulphuric acid and hydrochloric acid in water, in both of which hydrogen is liberated at the kathodes), the masses there set free are the same, the chemical equivalent of any one substance being always the same; but the quantities of electricity carried through each electrolyte are the same, and they are carried on the same number of ions in each, since the masses are equal; consequently the charge carried on an ion in the one electrolyte is the same as that carried on the same ion in the other electrolyte. There is, then, always associated with an ion of any substance a definite charge which cannot be varied; it is just as much a property of the ion as its mass.

Second, the charges carried on ions of different substances are in the ratio of their valences. This is proved by the following deduction from Faraday's laws: let the same current pass through two electrolytes in series, so that the same quantity of electricity is being carried by two sets of ions. A certain number of ions in each set are liberated in a certain time; call these numbers n_1 and n_2 for the two sets; and call the electric charges carried on each ion of the two sets e_1 and e_2. Then

$$n_1 e_1 = n_2 e_2,$$

since the same total charges are being carried. Further, if m_1 and m_2 are the masses of the two substances set free, and w_1 and w_2 the actual masses of an ion in each of the two sets,

$$m_1 = n_1 w_1,$$

$$m_2 = n_2 w_2;$$

that is,

$$m_1 : m_2 = w_1 e_2 : w_2 e_1.$$

But by Faraday's second law, writing c_1 and c_2 for the chemical equivalents of the two ions,

$$m_1 : m_2 = c_1 : c_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Substituting for m_1 / m_2 its value, this becomes

$$w_1 e_2 : w_2 e_1 = c_1 : c_2.$$

But $w_1 / w_2 =$ ratio of "atomic weights" by definition of "atomic weight;" and hence, since c is the ratio of atomic weight to valence, $e_1 / e_2 =$ ratio of the valences of the two substances. So, if a hydrogen ion carries a charge e, any other ion, whose valence is v, must carry a charge $v\,e$.

This fact follows at once if ions are regarded as dissociated parts of molecules. For a molecule is uncharged or neutral; and so, if it breaks up into ions, some will be $+$, the others $-$; but the amount of each charge must be the same. Thus, if a molecule of hydrochloric acid breaks up into an hydrogen ion and a chlorine ion,

$$HCl = H + Cl,$$

the charge on the hydrogen ion must be equal but opposite to that on the chlorine ion. If a molecule of zinc chloride breaks up into a zinc ion and two chlorine ions

$$ZnCl_2 = Zn + Cl + Cl,$$

the charge on the zinc ion must be twice as great as that on a chlorine ion, and therefore twice as great as that on a hydrogen ion. But the valence of zinc is two. All other ions may be considered in the same way; and the above law is a natural consequence.

If the electrolyte is a liquid, the same ion is always charged with the same kind of a charge; thus a hydrogen or a metal ion of any kind is always charged positively, an oxygen or chlorine ion always negatively. This is not true if the electrolyte is a gas; for there a hydrogen ion is under certain conditions positive, under others negative; and similarly with an oxygen or chlorine ion. The *amount* of the charge is, however, always the same for any one ion.

246. It is a matter of easy experiment to determine how many grams of any substance are liberated from an

electrolyte when a certain current passes; and, if this is known for any one substance, the amounts of any other substance which would be liberated by the same current may be calculated from Faraday's second law. The amounts which would be liberated by a different current may be calculated by Faraday's first law. The quantity of electricity, that is, the intensity of the current multiplied by the number of seconds ($i\,t$), would be naturally calculated in terms of the electrostatic unit (see Art. 218); but it is practically impossible to measure currents in terms of this unit. So another unit is ordinarily used for measuring currents; and, if its value is determined once for all in terms of the electrostatic unit, the current may be expressed in terms of it also, if it is desired. This new unit of quantity for measuring currents will be defined later (see Art. 277); and it is called the "electromagnetic unit." The number of grams of any substance liberated when 1 electromagnetic unit passes (i. e. the number of grams of a definite kind of ions, which carry a charge equal to 1 unit) is known as the "electro-chemical equivalent" of that substance. Some of these values for different elements are given in the following table: —

TABLE XIV

ELECTRO-CHEMICAL EQUIVALENTS

Chlorine	0.0003675	Iron (Ferric)	0.001932
Copper (Cupric)	0.003261	Oxygen	0.000828
Hydrogen	0.00010352	Silver	0.011180

It should be observed that, if a certain number of units of quantity liberate a number of grams of a substance equal to its chemical equivalent, that same quantity will liberate a number of grams of any other substance equal to its chemical equivalent. For, by Faraday's second law,

$$m_1 : m_2 = c_1 : c_2.$$

And, if $m_1 = c_1$, m_2 must equal c_2. It may be easily calculated from the table just given that 9654 units will set free an amount of any substance whose mass numerically equals its chemical equivalent. (The chemical equivalent of hydrogen is 1; hence 1 gram of hydrogen ions carry 9654 electromagnetic units.)

247. Dissociation Theory of Electrolytes. As already stated, the laws of electrolysis may all be explained if the ions are dissociated portions of molecules, each one carrying a definite charge proportional to its valence. This theory will now be explained more in detail.

In gases the explanation is extremely simple. It is known that gases can be dissociated in several ways, by raising the temperature sufficiently, by producing so great an electric strain that a spark passes, etc.; and experiments prove that, if there is any dissociation in a gas, the gas can conduct a current, and, if there is no dissociation, it cannot.

In the case of liquid electrolytes, the explanation is more complicated. It must first be noted that, with the exception of certain fused solids, no liquid can conduct a current unless there is dissolved in it some salt or acid. Further, only particular kinds of liquid solutions can conduct; and in all these, with no exception, there are independent reasons for believing that the salt or acid is dissociated. The theory is that, when a salt or acid is dissolved in certain liquids, there is partial dissociation. This is particularly true of ordinary salt or acid solutions in water, but is not true of most organic solutions, such as sugar in water or salt in benzine. If, for example, a gram of common salt, sodium chloride ($NaCl$), is put in water, a definite percentage of the salt will dissociate into sodium and chlorine ions. These ions are not molecules, but are charged atoms in this case; and so it would not be expected that the usual properties of solid sodium or chlorine gas would be shown. The sodium ion is positively charged; the chlo-

rine ion negatively, as is proved by passing a current through the liquid, and seeing which substance comes with the current to the kathode, and which against the current to the anode. In this solution of common salt in water the percentage of molecules dissociated is supposed to remain constant (unless temperature or other conditions change); but the process of dissociation and combination is supposed to go on continually, so that any individual ion may exist as such for a while, and may then combine with an ion of the other kind to form a molecule. This process is supposed to be going on in any solution of common salt in water. If, now, a difference of potential is applied at two points in the solution by two wires joined to a primary cell (or any other source of E. M. F.), there will be an electric force urging the positively charged ions towards the kathode during those intervals of time in which these ions are dissociated; similarly, the negatively charged ions will be urged in the opposite direction, towards the anode. Any individual ion will be given a velocity in one direction, may then meet another ion of the other kind, with which it combines, and will then receive no further acceleration until it becomes free again. It must be noticed that, on this theory, the dissociation is not produced by the E. M. F., or the current, but takes place as a part of the process of solution.

In any actual case it is exceedingly difficult to determine exactly what the ions are. If pure water is dissociated at all, it must be to an amount almost too minute to measure. Any substance which goes into solution and dissociates forms ions; and these are carried to the anode and kathode. The actions at these places, however, are complicated. Some ions combine with each other to form molecules; others react on the liquid, and the products of the reaction form molecules and are liberated. Of course no substance can be liberated unless it exists in the form of molecules. Consider several cases.

1. Dilute aqueous solution of sulphuric acid, H_2SO_4, with platinum wires as anode and kathode. The ions are supposed to be

$$\overset{+}{H}, \overset{+}{H}, \overset{=}{SO_4}.$$

The H ions go to the kathode, give their charges to the metal conductor there, combine to form molecules, H_2, and so bubble off as a gas. The SO_4 ions go to the anode, give up their charges to the metal conductor, but cannot escape, because of themselves they cannot form molecules, neither can they combine with the platinum wire; so there is a reaction with the water,

$$2\,H_2O + 2\,SO_4 = 2\,H_2SO_4 + O_2,$$

and the oxygen gas bubbles off.

2. A solution of common salt, NaCl, in water, with platinum wires for anode and kathode. The ions are supposed to be $\overset{+}{Na}$, $\overset{-}{Cl}$. The Na ions go to the kathode; and, in general, there is a reaction with the water.

$$2\,H_2O + 2\,Na = 2\,NaOH + H_2;$$

so hydrogen gas escapes. Chlorine gas escapes at the anode.

3. A solution of copper sulphate, $CuSO_4$ in water. The ions are $\overset{++}{Cu}$, $\overset{=}{SO_4}$. The Cu ions go to the kathode, where they will in general be deposited as metallic molecules. The SO_4 ions will go to the anode; and, if this is a platinum wire, oxygen gas will be evolved as in case 1. But, if the anode is a copper plate, the SO_4 can react on the copper, making some of it go into solution. So that, in this last case, copper is dissolved from the anode and deposited on the kathode. This is the principle of "copper-plating." Silver-plating may be done in a similar manner by using a silver solution which is a conductor and a silver plate as an anode.

248. As stated above (Art. 242), when a solid conductor dips in a liquid one, there is always a difference of

potential between the two. This fact may be explained at once on the dissociation theory of electrolytes. The simplest case is when a metal dips into a solution of one of its own salts, e. g. zinc into a solution of zinc chloride ($ZnCl_2$). When the solution of $ZnCl_2$ is made, it is supposed that some of the molecules dissociate, thus giving $\overset{+}{Zn}$ and $\overset{-}{Cl}$ ions. If now a rod of zinc dips in the solution, some of the zinc in general dissolves (unless the solution is "supersaturated"); and, since experiments show that in a liquid solution of any metallic salt, the metal ion is always positively charged, the zinc ion which leaves the rod and enters the solution has a positive charge. But there must be an equal negative charge left on the zinc rod; and owing to the action of positive and negative charges, the positive charged zinc ions in the solution will form a layer around the negatively charged zinc rod. Consequently there will be a difference of potential between these two layers, the zinc rod being lower than the solution. As the process continues, the electric energy of these two layers increases rapidly; and, when any further solution of the zinc rod would produce an increase in the entire potential energy of the various changes, the process stops. This is the reason why a pure metal cannot perceptibly dissolve in an acid; an electric current is required. Further, since a metal ion leaves the metal carrying with it a positive charge, it is seen why in a circuit carrying an electric current, that metal dissolves at whose surface the current flows *into* the electrolyte, if there is any chemical action between the solid and liquid; while at the other metal pole, the kathode, a metal may be deposited.

On the dissociation theory, no charges of any kind can exist except on portions of molecules, i. e. ions. A charged conductor has, then, on its surface a great number of dissociated molecules; and the process of charging a body consists in causing dissociation of its surface molecules into positive and negative ions, and then making the

charges all the same, either by removing those ions which are charged the opposite way or by reversing their charges.

249. Process of Conduction. The process of conduction in an electrolyte consists, then, in the handing on of the positive charges in one direction, and of the negative ones in the other, both charges being carried by portions of matter called ions. This is the mode of action in most liquid and gaseous conductors. In the case of solid metal conductors the process may be just the same, there being no reason for believing it different in principle. Of course, the molecules of a solid cannot travel far, — they can simply oscillate about positions of equilibrium; and, again, only a very small proportion of the molecules need take part in the conduction. In all cases, though, the process is essentially discontinuous.

250. Discharge through Gases. When an electric discharge, or spark, passes through a gas, there are many most interesting and important phenomena. The simplest illustration of such a discharge is given by a flash of lightning. The methods commonly used for producing a discharge are to cause sparks to pass by means of an electric machine, such as a Voss or a Wimshurst one, or by means of a so-called induction-coil (see Art. 289). The gas through which the discharge takes place is rendered luminous by its passage; and the light observed is characteristic of the vapor or gas, and also of the metal poles between which the spark passes, because in the act of discharge some of the metal molecules are torn off from the terminals of the machine or coil.

The nature of the discharge depends largely upon the pressure of the gas. If it is at ordinary pressures, the general appearance is like that of the lightning; but, if the gas is enclosed in a vessel from which some may be exhausted by means of an air-pump, it is found that at low pressures the character of the discharge entirely changes. The method of observation generally adopted is to place

the gas in a glass bulb, into which enter metal wires; and to study the effect of passing discharges between these metal electrodes, at different pressures of the gas.

If the pressure of the gas is in the neighborhood of 1 mm. of mercury, the typical discharge is like that shown in the diagram. At the anode there is a series of stratifications or "striæ," generally beautifully colored. While at the kathode there is a dark space except immediately on the surface of the metal pole. These two phenomena are entirely distinct and independent of each other. It has been recently shown that the existence of striæ in a dis-

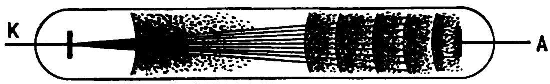


FIG. 175.

charge-tube is probably positive proof of the presence of an impurity in the gas. There is also another phenomenon at the kathode, which deserves notice, and which is best observed if the vacuum is such that the pressure of the gas is only a fraction of a millimetre of mercury, i. e. if the gas forms the "Fourth State of Matter" (see Art. 210). In a gas so rarefied, the phenomena at the anode disappear, the dark space from the kathode fills the tube; but now it may be plainly observed that there is a faint radiation from the face of the metal forming the kathode, and that it takes place in straight lines. This may be shown by the faint luminescence of the gas through which the rays pass, and by the brilliant colored effects produced where the rays strike the glass walls of the tube or other substances which emit light under these conditions. By putting obstacles between the kathode and the walls of the tube, it may be shown that the rays pass in straight lines, because such sharp shadows are cast. These rays are called the "kathode rays;" and they have been proved to consist of streams of portions of matter, negatively charged, moving

away from the kathode with a great velocity. These rays can pass through many obstacles if they are thin enough, such as aluminium foil; they can affect a photographic plate; they are deflected by a magnet (like any electric current).

Besides these two radiations produced in a highly rarefied gas, viz. light waves and "kathode rays," there is another which is not yet understood. When the discharge is taking place through the tube, it may be proved that, proceeding out from some points inside, through the glass, there is a distinctly new radiation. This was discovered by Professor Röntgen of Würzburg, Germany; and many of its properties are now known. The rays proceed in straight lines; they are not deflected by a magnet; they make many substances emit light when they fall upon them; they can pass through many obstacles which are opaque to ordinary light waves; they affect photographic plates; they change any insulating medium or dielectric, such as air, paraffin, etc., into a conductor when they pass through it.

CHAPTER IV

PROPERTIES OF STEADY ELECTRIC CURRENTS

WHEN an electric current is produced by applying a constant difference of potential, or E. M. F., to a conductor or series of conductors, certain effects are noticed in the conductors, and these take place according to certain laws. Such a current is called a "steady" current.

251. Uniformity of Current. It may be proved by most careful experiments that the intensity of a steady current is the same in all parts of the circuit, no matter what the conductor is,—liquid, gas, or solid. And, if the circuit branches at any point, so that two or more conductors carry the current away from that point, the sum of the intensities in these conductors exactly equals the intensity of the current in the conductor by which the current is brought to the point of branching. If this law were not true, there would be an increasing or decreasing statical charge at certain points in the circuit; and this would show that the current was not steady. Further, it may be proved directly by showing that the magnetic effect of the current (see Art. 253) is the same at all points of the circuit.

252. Heating Effect. The current always flows from high potential to low; and in so doing energy is given the minute portions of matter which make up the conductor; electric energy disappears, being transformed into heat energy. Stated in other words, work is required to make a current pass along a conductor; and, this work being done against forces connected with the smallest portions

of the conductor, heat effects are produced. (Notice the analogy to water flowing from a point of high pressure to one of low through a pipe where there is a good deal of friction.) The fact that the temperature of a conductor rises when a current passes in it, is familiar to every one by a great many illustrations, such as the ordinary glow-lamp, the arc-lamp, the burning out of dynamos and motors, the violent sparks produced wherever there is a bad contact, etc.

The work done in any particular portion of the conducting circuit may be easily calculated. Let A and B be two points on any conductor forming part of the circuit through which a current of intensity, i, flows. Let the current flow from A to B, and let E be the difference of potential between them.

FIG. 176.

Therefore, in t seconds a quantity $i\,t$ passes, because i is the quantity that passes in 1 second. But E is the amount of work *done by* a unit positive quantity of electricity in passing from A to B, or by a unit negative quantity in passing from B to A. So in t seconds the work done by the current against these minute forces of the conductor is $E\,i\,t$. This is generally called the heating effect, and may be written

$$H = E\,i\,t \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

H is, of course, measured in ergs. The rise in temperature may also be calculated. If m is the mass of the conductor between A and B; c, its specific heat; T, the rise in temperature; J, the mechanical equivalent of heat (see Art. 212), 4.2×10^7; then

$$H = m\,c\,T\,J = E\,i\,t;$$

or

$$T = \frac{E\,i\,t}{m\,c\,J} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

253. Magnetic Effect. If a magnetic needle is suspended, free to turn, near a current, the magnet is affected. In particular, if the current is horizontal, and the magnetic needle is suspended over it and parallel to it, the magnet is deflected so as to stand as nearly as it can at right angles to the current. The fundamental property of a magnetic needle is, of course, that it tends to point in a direction which is nearly north and south; and that end which tends to be towards the north is called the "north pole" of the magnet; and the other end, the "south pole." It may be proved by experiment that the action of a current on a magnet may be thus described: the north pole of a magnetic needle tends to move around a current in such a direction that there is the same connection between the direction of the current and this direction of the motion of the north pole of the magnet as there is between the direction in which an ordinary right-handed screw advances into a board and the direction in which it must be turned in order to make it enter. This law of direction is sometimes called "the right-handed screw law." The south pole of the magnetic needle tends to move in an opposite direction. Consequently, a magnetic needle will tend to place itself at right angles to the current. In particular, if a magnet is suspended inside a coil of wire, parallel to the coils, it will be deflected when there is a current through the wire. Let the direction of the current and the position of the needle be as shown. The current *over* the magnet will tend to drive the north pole down into the plane of the coils, the south pole up out of the plane, as shown. Similarly, the current below the

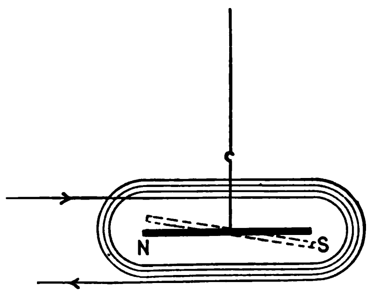


FIG. 177.

where ρ is a quantity depending only upon the material of the conductor. If the temperature changes, ρ changes, too; it increases for most solid conductors, if the temperature is raised; and decreases for most liquid conductors.

The reciprocal of the resistance is called the conductance, C. Its value is, then,

$$C = \frac{1}{R} = \frac{\sigma}{\rho l} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (5)$$

ρ is sometimes called the "resistivity" or "specific resistance," and $\frac{1}{\rho}$ is also called the "conductivity" or "specific conductance."

TABLE XV

SPECIFIC CONDUCTIVITY, REFERRED TO MERCURY

Aluminium (soft)	32.35	Nickel (soft)	3.14
Copper (pure)	59.	Platinum	14.4
Iron	9.75	Silver (soft)	62.6
Mercury	1.000	Tin	7.

RESISTANCE, IN OHMS, OF WIRE 100 CM. LONG, 1 MM. CROSS-SECTION

		Rate of Change in Resistance per degree Centigrade.
Aluminium	0.02916	0.00388
Copper	0.0160	0.00380
German Silver	0.2670	0.0035
Iron	0.0973	0.00650
Mercury	0.9434	0.000907
Platinum	0.0907	——
Silver	0.01506	0.00377

These values in terms of "Ohms" are not exactly in the unit defined later in Article 280. The Ohm here used

is a standard resistance which differs extremely slightly from the true Ohm as defined later.

Applications of Ohm's Law.

255. *a.* **Conductors in Series.** If several conductors, AF, BC, CD, are joined in series, Ohm's law may be applied to each one separately. Let R_1, R_2, R_3, be the resistances of the conductors, and V_A, V_B, etc., the potentials at A, B, etc. Then, since the intensity of the current is the same throughout (see Art. 251),

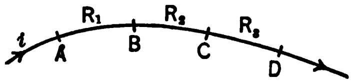


Fig. 181.

$$i = \frac{V_A - V_B}{R_1} = \frac{V_B - V_C}{R_2} = \frac{V_C - V_D}{R_3}.$$

Hence, if there is no E. M. F. at the junctions A, B, C,

$$i = \frac{V_A - V_D}{R_1 + R_2 + R_3}.$$

Therefore, the entire resistance between A and D is

$$R = R_1 + R_2 + R_3 \; . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

That is, when several conductors are joined in series, their entire resistance is the sum of the resistances of the separate conductors. (This shows why in a uniform conductor R is proportional to its length.)

256. *b.* **Conductors in Parallel.** If a conductor branches

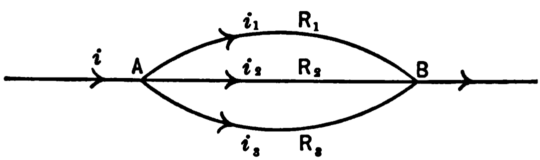


Fig. 182.

so that several conductors carry the current between two points, they are said to be in parallel. Thus, let three

conductors, of resistances R_1, R_2, R_3, be joined in parallel between A and B; and let the currents in them be i_1, i_2, i_3. The total current

$$i = i_1 + i_2 + i_3,$$

and the entire resistance between A and B is R in the formula

$$i = \frac{V_A - V_B}{R}.$$

But
$$i_1 = \frac{V_A - V_B}{R_1},$$
$$i_2 = \frac{V_A - V_B}{R_2},$$
$$i_3 = \frac{V_A - V_B}{R_3}.$$

Hence
$$i = (V_A - V_B)\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right),$$

or
$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \quad \ldots\ldots\ldots \quad (7)$$

Expressed in terms of conductances, the total conductance

$$C = C_1 + C_2 + C_3.$$

That is, when several conductors are joined in parallel, the total conductance equals the sum of the conductances of each conductor. (This shows why R varies inversely as the cross-section of a conductor.)

Comparing the values of i_1, i_2, and i_3, it is seen that

$$\left.\begin{aligned} i_1 : i_2 &= \frac{1}{R_1} : \frac{1}{R_2} \\ i_2 : i_3 &= \frac{1}{R_2} : \frac{1}{R_3} \end{aligned}\right\} \quad \ldots\ldots \quad (8)$$

That is, when a current is divided among several branches, the intensities in any two branches are in the inverse ratio of their resistances.

257. *c.* "**Wheatstone's Bridge.**" This is a name given to a particular arrangement of conductors, which consists, in principle, of two conductors, $A\,B\,D$ and $A\,C\,D$, in parallel, having two points, B and C, joined by a third conductor, and having a fourth conductor joining A and D directly. If a primary cell is placed in this last conductor, $A\,D$, a cur-

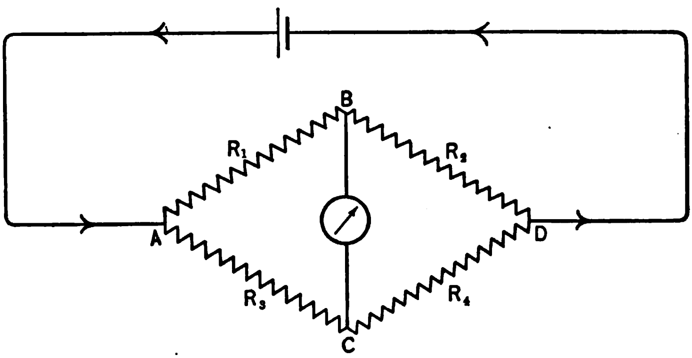


Fig. 183.

rent will flow either to A or to D, and will be divided there into two branches. Call the resistance of the conductor $A\,B$, R_1; of $B\,D$, R_2; of $A\,C$, R_3; of $C\,D$, R_4; and let the respective current intensities be i_1, i_2, i_3, i_4.

Then, by Ohm's law,

$$i_1 = \frac{V_A - V_B}{R_1}; \quad i_2 = \frac{V_B - V_D}{R_2},$$

$$i_3 = \frac{V_A - V_C}{R_3}; \quad i_4 = \frac{V_C - V_D}{R_4}.$$

There will in general be a current through the conductor which leads from B to C; but, if there is none, i. e. if $V_B = V_C$, then $i_1 = i_2$, and $i_3 = i_4$. And in this particular case it follows from the above equations that

$$\frac{R_1}{R_3} = \frac{R_2}{R_4},$$

or

$$R_1 R_4 = R_2 R_3 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

It is perfectly easy, though, to secure the condition that $V_B = V_C$. If a galvanoscope is placed in the conductor leading from B to C, a current may be instantly detected; and if the point C (or B) is shifted along the conductor $A\ C\ D$ (or $A\ B\ D$), it must be possible to find a point such that no current flows through the galvanoscope. Then, if R_3 and R_4 are the resistances between this point C and A and D, respectively, the above equation is true.

Consequently, if the numerical values of three resistances are known, that of the fourth may be determined. The unit of resistance will be defined later (see Art. 279); but Wheatstone's bridge gives at once a method of comparing resistances.

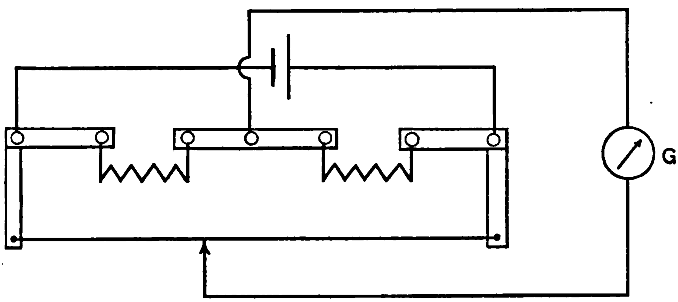


FIG. 184.

A simple modification of the bridge is to make the conductor, $A\,C\,D$, a long uniform wire. In which case, if l_3 and l_4 are the *lengths* of the wire from A to C and from C to D,

$$R_3 : R_4 = l_3 : l_4;$$

and so, when no current flows from C to B,

$$R_1 : R_2 = l_3 : l_4 \quad . \quad . \quad . \quad . \quad . \quad (10)$$

The lengths l_3 and l_4 may be measured; and thus the ratio of any two resistances may be easily determined. Fuller details of the method are given in all laboratory manuals;

and so are also other methods for the comparison of resistances.

258. *d.* **Heating Effect.** As proved in Article 252, the heating effect in any portion of a conductor through which a current of intensity i flows for t seconds is $E\,i\,t$, if E is the difference of potential at the two ends of the conductor. By Ohm's law, if the current is steady, $E = i\,R$. Hence, substituting for E in equation (1) its value,

$$H = i^2 R\, t \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

This shows that, if a current is flowing through any series of conductors, the heat effect is greatest where the resistance is greatest. This explains the great rise in temperature wherever there is a bad contact in the circuit.

This shows, too, that the heating effect through a conductor is independent of the direction of the current; for, if the current is reversed, the intensity becomes $-i$, and $H = i^2\, R\, t$, as before.

This same formula may be expressed also in terms of E and R; for, substituting $i = E\,/\,R$, it becomes

$$H = E^2\, t\,/\,R \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11\,a)$$

CHAPTER V

GENERAL PROPERTIES OF MAGNETS AND MAGNETIC FIELDS

IT is well known to every one that an ordinary magnet has the power of attracting pieces of iron or steel; and that, if a light magnetic needle is pivoted so as to be free to turn, it will tend to turn into a direction nearly north and south. It is also probably known to all that there is an intimate connection between an electric current and a magnet, as is illustrated in the last chapter (Art. 253). Before discussing this last property of a magnet, it will be necessary to consider in detail the first two.

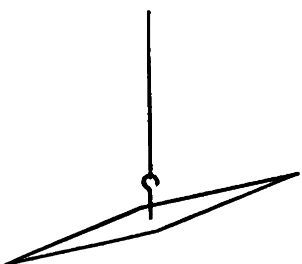


FIG. 185.

259. Definitions. A magnet may be defined as a body which, without being electrified, has the power of attracting pieces of iron in its natural state. This means that work is required to move a piece of iron away from a magnet, and that consequently the potential energy will be decreased if the iron and the magnet approach each other. So they tend to approach, and will do so unless hindered by some external force. This energy which is associated with magnets is in the ether; and the presence of ordinary matter influences the phenomenon, — for example, the forces in air are entirely different from those in pure oxygen. But when magnets are near each other, there is

no such statical mechanical strain of the surrounding medium as there is in the case of charged bodies. In fact, as will be more clearly seen further on, the phenomena of magnetism are evidently conditioned by *motion* in the ether, and in the smallest portions of the surrounding medium.

Other substances than iron are attracted by magnets when the surrounding medium is air; among them are nickel, cobalt, manganese, and many compounds of iron. Such bodies are called "magnetic substances."

Other bodies are repelled by magnets, when the surrounding medium is air; such are bismuth, antimony, and zinc. All such bodies are called "diamagnetic substances;" but the forces of repulsion are extremely minute compared with the forces of attraction of a magnet for iron.

Attraction or repulsion will, of course, take place, so that the potential energy may decrease. (See Art. 221 for the similar case in electricity.) The motion in air may, therefore, be different from that in other media; the question is one simply of relative amounts of potential energy in equal volumes of different media.

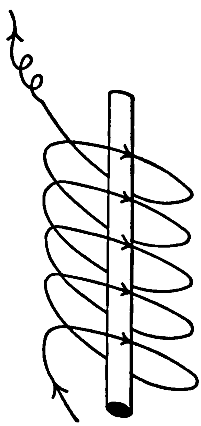


FIG. 186.

260. Magnets. There is one natural substance which is a magnet, viz. the so-called magnetic oxide of iron (Fe_3O_4), which occurs as an ore in many parts of the world. But it is produced in lumps, and so is not generally useful. Any magnetic substance, such as steel, iron, etc., may, however, be made a magnet; and convenient shapes, such as cylinders, bars, and needles, may be chosen. The simplest way of making a bar of iron or steel a magnet is to place it in a helix or spiral coil of wire through which an electric current is passing. (The current passes around

the bar of iron, *not* through it.) If this is done, the bar is found to be a magnet, and remains so when removed from the current. If the material is soft iron, the magnetism will disappear, however, at the least disturbance, such as tapping it; whereas, if the bar is steel, it will remain a magnet for a long time, and is sometimes called a "permanent" magnet. (The fuller explanation of this method of making a magnet will be given later, in Chapter VII.) If such a permanent magnet is once secured, it can be used to make other magnets; for, if this magnet is drawn lengthwise over a bar of iron or steel, the latter itself becomes a magnet. If a bar of iron or steel is even placed near a permanent magnet, it becomes itself a magnet, although a weak one.

261. Polarity. If a bar or needle of steel is magnetized by being placed along the axis of a helix of wire carrying a current, it is said to be magnetized "lengthwise;" and if, after being removed from the helix, it is pivoted or suspended free to turn, it will place itself in a direction nearly north and south, in what is called the "magnetic meridian." This shows that the two ends of the magnet have opposite properties: one is attracted towards the north, the other is repelled from the north. The bar or needle is said to be "polarized;" and the end towards the north is called the "north pole" of the magnet; that towards the south, the "south pole."

If the action of one bar-magnet on another is examined, it is seen that two like poles repel each other, while two unlike poles attract each other. (It must not be thought that all magnets have poles; it is true only of bar-magnets and needles, in general.) The exact position of the "pole," considered as a point, is not definite. If the forces in the neighborhood of a bar-magnet are examined by means of a small magnetic needle, it may be proved that there are magnetic forces at all points of the bar-magnet; so that, if the bar-magnet is placed at right

angles to the magnetic meridian, there is a series of parallel forces attracting one half the magnet towards the north, and another series of parallel forces repelling the other half. The resultants of each of these two sets of parallel forces are two parallel forces; and they meet the magnet in two points which may be called its "poles." The distance between these two poles is called the "length" of the magnet.

262. Equality of Poles. If a bar-magnet is placed on a cork which floats in water, so that it is free to move in any way, it is observed that the magnet does not move as a whole in any direction. Its centre of inertia remains fixed; and it turns around an axis through this until it comes to rest in the magnetic meridian. This proves that the forces on the two poles are exactly equal and opposite; or, in other words, the poles are of equal "strengths."

If any other magnet, even a lump of magnetic ore, is placed on the cork, it also will have no motion of translation, showing that all the forces due to the magnetic field of the earth reduce to a couple.

263. Molecular Nature of Magnetism. There is every reason for believing that magnetism is a molecular property of a body; that every molecule of a magnet is itself a magnet. For, anything which influences the molecules of a magnet affects the magnetism as well. Thus, increase in temperature decreases the magnetism; and at "red-heat" a steel magnet loses practically all its magnetism. Hammering a magnet will always alter its magnetism. Further, if a magnet is broken in two, each portion is found to be a magnet, no matter how minute the fragments are.

Since any magnetic substance (iron, steel, nickel, etc.) may be made a magnet, the entire phenomena of magnetization may be explained if each molecule of a *magnetic substance* is a magnet with equal poles. In a bar of such a substance the molecules may be considered as lying at random, and therefore neutralizing each other's external

action; and the process of making such a bar a magnet consists simply in turning the molecular magnets into more or less the same direction. For, if this is done, the north poles being in one direction and the south poles in the other, there will be a resultant action near the two ends of the bar. (The molecules in the figure are drawn elongated, simply to liken them to ordinary bar-magnets, not to

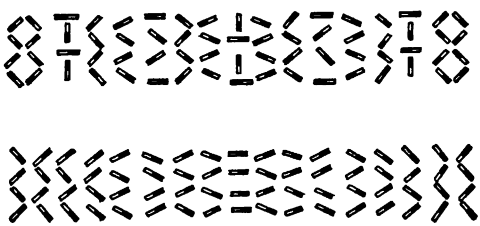


FIG. 187.

imply that a molecule is like that.) This explains also why the two poles of a bar-magnet are equal, because each molecular magnet has equal poles. Similarly, even if the magnet is not in the shape of a bar or needle, but has minute poles all over it, it is seen that the north poles must equal the south poles in strength.

It is evident, too, why a bar of iron becomes a magnet when a permanent magnet is rubbed over it lengthwise or even placed near it, because that pole of the permanent magnet which passes over the iron, or which is nearest it, attracts all the opposite poles of the molecular magnets of the iron, and repels all the like poles, thus turning the molecules into a more or less parallel direction. If they are all forced into a strictly parallel direction, the magnetization must be the greatest possible. It is observed, in

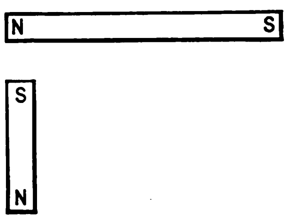


FIG. 188.

fact, that a bar of iron or any magnetic substance cannot be magnetized beyond a definite limit, in which state it is said to be "saturated."

Another evidence in proof of the fact that each molecule of a magnetic substance is a magnet is the fact that it is possible to make a model of such a substance which exhibits all its properties. If an immense number of small pivoted magnets are placed near each other on a table, they will behave exactly as do the molecules of a magnetic substance. At first they will stand perfectly at random; then, if a current is passed around them by placing the table in a helix of wire, or if a strong magnet is brought near, the small magnets will place themselves more or less parallel; and, if the current is stopped or the magnet removed, the small magnets will still remain nearly parallel, unless disturbed by jarring.

A further evidence is afforded by the fact that when a bar of a magnetic substance is magnetized suddenly, a distinct click is heard, known as the "Page effect;" and the length and volume of the bar also change (except in quite special cases).

Since, then, magnetism is a molecular property, and each molecule of a magnetic substance has both a north pole and a south one, it is evident that a north pole cannot be separated from its equal south pole; the same magnet must always have both north and south poles. (This is unlike electrical charges, which can be separated so as to be on different bodies.)

264. "Unit Poles." Two equal poles may, of course, be defined as two poles which have identically the same effects, e. g. exert the same forces on a suspended magnet. And, if two equal poles when placed one centimetre apart in air, exert a force on each other of one dyne, they are each said to be "unit poles;" or each pole is said to have a "strength" one.

Law of Force. The law of action between magnets

may be stated in this way: if a pole of strength m is placed at a distance of r cm. from another pole of strength m' in any medium, the force in dynes is given by the formula

$$F = \frac{m\,m'}{\mu\,r^2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where μ is a quantity which depends upon the medium. It is not, in general, constant for any one medium, but varies with the amount of the magnetic force. It has received the name "permeability," for reasons which will be given later. This law of force is amply verified by experiment, because all deductions from it are in accordance with observed phenomena.

For air μ is constant; and it equals 1, on the above definition of a unit pole. For, if $m = m' = 1$, and $r = 1$ in air, F must equal 1; i. e. $\mu = 1$ for air. If a north pole has a strength m, an equal south pole must have a strength $-m$; because the force on a north pole is exactly opposite to that on an equal south pole.

For all magnetic substances μ is greater than 1, on the magnetic system of units above defined; and for all diamagnetic substances μ is less than 1. It should be noticed, though, that, since all diamagnetic forces are so feeble, μ may in practice be regarded as being equal to 1 except in magnetic media. These facts may be proved from analogy with similar facts in electricity, as follows.

265. Attraction and Repulsion. K, the dielectric constant in electricity, plays the same part in electrical phenomena that μ, the magnetic permeability, does in magnetic. But it was proved in Chapter I. Article 221, that, if any piece of dielectric, e. g. a bit of glass or sulphur, is attracted by a charged body in air (where $K = 1$), this is evidence that K for that dielectric is greater than for air; whereas, if a piece of dielectric is repelled, K for it is less than for air. But in magnetic phenomena magnetic bodies are attracted

in air by a magnet, and diamagnetic bodies are repelled. Consequently, μ for the former is greater than for air, and for the latter is less.

266. Magnetic Lines of Force. A magnetic field of force may be defined as the region in which magnetic forces are manifest; and magnetic lines of force may be drawn just as electric lines of force were, viz. a magnetic line of force gives at each point the direction in which a north pole would move if placed at that point. A south pole would, of course, tend to move in the opposite direction; and, consequently, if a very small magnet is placed at any point in a magnetic field, it will place itself tangent to the line of force at that point. (In this way the lines of force of any magnetic field may be mapped.) If the lines of force are all parallel, the field is said to be "uniform."

Lines of magnetic force always pass from north poles to south poles; and, since a magnet is made up of small mag-

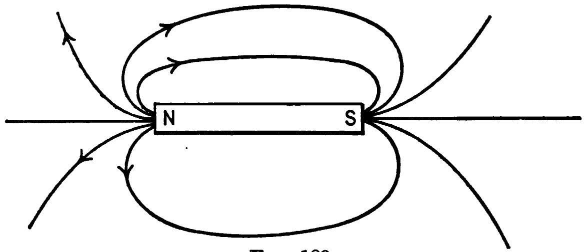


FIG. 189.

nets, lines of force pass *through* magnets, and do not end on their surfaces. A drawing is given of the lines of force around a bar magnet. Magnetic "potential" at a point in the field may be defined as the work required to bring up a unit north pole from infinity to that point. (The earth cannot in this case be taken as the standard, because its magnetic potential is not the same at all points.) So equipotential surfaces for magnetism may be drawn; and lines of magnetic force are always perpendicular to them, and pass from high to low potential.

The "intensity" of the magnetic force at any point of the field is the name given to the force which would act on a unit north pole if placed at that point; and wherever the lines of force are thickest the intensity is greatest, as is shown by the lines around a magnet.

267. Magnetic Induction. When any magnetic substance is placed near a magnet, i. e. is placed in a magnetic field, two phenomena are observed, as already noted: it becomes a magnet, and it is attracted by the permanent magnet. These phenomena have been already explained as consequences of the fact that the molecules of a magnetic substance are themselves magnets; for, if the north pole of the permanent magnet is nearer the magnetic substance than the south pole, the south poles of the molecular magnets will be turned towards it; and so there will be attraction. This phenomenon is known as magnetic "induction." The

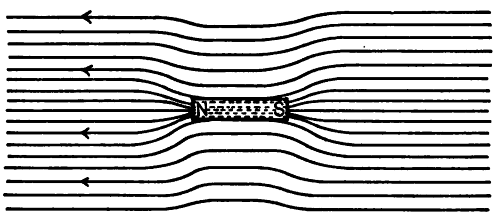


FIG. 190.

simplest case is when a bar of some magnetic substance is placed lengthwise in a uniform field of force in air. The field of force is found to be changed as shown. The lines of force crowd down into the magnetic substance, out of the air. (Compare Art. 229 and Fig. 158 for the similar case in electricity.) This phenomenon may be described by saying that the magnetic substance is more permeable for lines of magnetic force than air; and, since μ is greater for it than for air, μ is called the permeability. This change in the lines of force is exactly the same as would take place if there were superimposed upon the original

uniform field that due to a permanent magnet of suitable strength and of the same size and shape as the bar of magnetic material, with its south pole placed at the end into which enter the lines of force due to the uniform field. For, at any point, P, there would be two forces, as shown,—

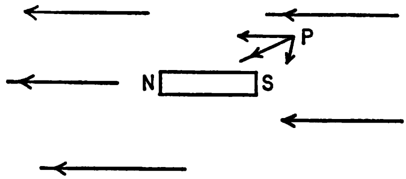


FIG. 191.

one due to the uniform field, the other to the magnet; and their resultant has the same direction as the line of force actually observed when the bar of magnetic material is placed in the uniform field. If the field of force is uniform inside the bar, as it will be near its middle point if the bar is very long, a small thin cavity may be imagined cut across the bar, or the bar may be imagined cut in two and the parts separated slightly; and the intensity of the magnetic force in that cavity is called the "magnetic induction." Its ratio to the intensity of the magnetic force in the original uniform field in air may be proved to be the value of μ for the magnetic substance.

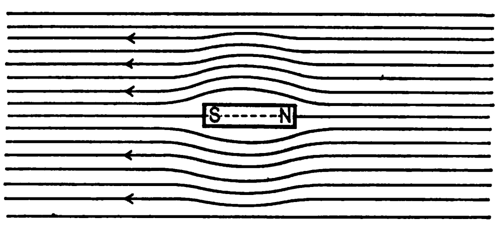


FIG. 192.

Similarly, if a diamagnetic substance is placed in a uniform field, the modification in the lines of force is as

shown; the lines crowd out into the air; the diamagnetic substance is less permeable. This change is exactly the same as would take place if there was superimposed upon

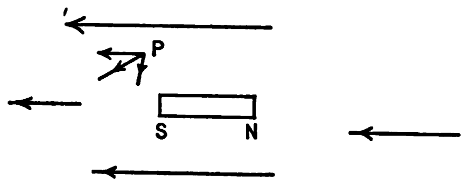


FIG. 198.

the original uniform field that due to a magnet of suitable strength and of the same size and shape as the diamagnetic substance, but with its north pole at the end into which enter the lines of force of the uniform field. This is shown in the drawing.

Of course the attractions and repulsions take place in accordance with the general law that the potential energy tends to decrease. If a magnetic substance is placed anywhere in the field, there will be less energy in it than in an equal volume of air at that same point in the field; and, the more intense the field, the greater the difference in the energy between the magnetic substance and the air. Consequently, a magnetic substance will move into that portion of the field where the intensity is greatest, while a diamagnetic substance will move into that portion where the field is weakest. In a uniform field, neither would move.

Since a magnetic substance is more permeable to lines of force than air, if a hollow sphere of iron is placed in a field of force, there will be comparatively no lines of force inside, because nearly all will pass through the iron shell. Similarly, if an iron ring is placed in a uniform field, the lines of force will be as shown, the field being most intense at the top and bottom of the ring.

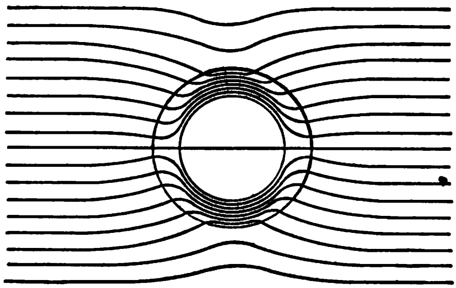


Fig. 194.

268. Magnetic Moment. If any magnet is placed at random in a magnetic field, and is suspended so as to be free to rotate around the axis of suspension, it will in general turn and tend to place itself in some particular position, i. e. there is a mechanical moment acting on it due to the field of force. The simplest case is when the field is uniform, and the intensity at each point is 1. Such a field is called a " unit field." The greatest value that the mechanical moment can have, which acts on the magnet when placed in any position in this unit field, is called the " magnetic moment " of the magnet. The simplest magnet is a bar-magnet with its two poles of strength, m, at a distance, l, apart. The maximum moment which it can experience in a unit field is obviously that which it has when it is at right angles to the field. There will then be a force of m dynes, parallel to the field, acting on its north pole, because by definition of intensity it is the force acting on a unit pole; and, if the intensity is 1, the force on a pole of strength m is m dynes. Similarly, there will be a force $-m$ dynes acting on its south pole; i. e. a force exactly opposite to the one on the north pole. These two forces form a couple; and, since the distance between the

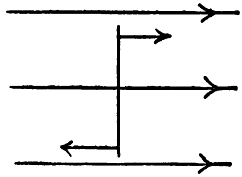


Fig. 195.

poles is l, the moment of the couple is $m\,l$. If the magnetic moment is written M, then for a bar-magnet

$$M = m\,l \;.\;.\;.\;.\;.\;.\;.\;.\;. \quad (2)$$

If a bar-magnet is placed at an angle θ with the direction of the unit field of force, the moment acting on it is $M \sin \theta$, because the distance between the two parallel forces is now $l \sin \theta$. If the uniform field of force has an intensity R, the force acting on the pole m is $m\,R$ dynes; and so when the bar-magnet is placed at an angle θ with the direction of such a field, the mechanical moment is $R\,M \sin \theta$.

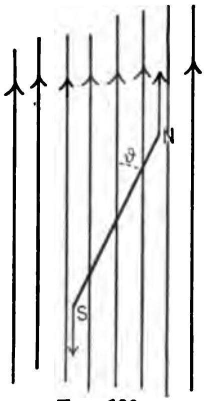


FIG. 196.

269. Measurement of R M. When a magnet is not parallel to the field of force, there is a couple acting on it tending to make it turn into that direction. So, if a magnet is suspended free to turn about an axis perpendicular to the field, and is displaced slightly out of the direction of the field of force, it will be turned back into that direction, and will swing through it, then return, etc., making oscillations until it is brought to rest by friction. The mechanical moment at any instant is $R\,M \sin \theta$; or, if the angle is very small, $R\,M\,\theta$. This moment is proportional to θ, the angular displacement; and it tends to bring the magnet back into the direction of the field of force; consequently, the motion is simple harmonic vibration (see Art. 25); and, if A is the moment of inertia of the magnet around the pivot as an axis, the period of one complete vibration is

$$T = 2\,\pi \sqrt{\frac{A}{R M}}, \;.\;.\;.\;.\;.\;. \quad (3)$$

for the angular acceleration is $\dfrac{R\,M\,\theta}{A}$, the moment of the forces divided by the moment of inertia.

Consequently $$R M = \frac{4 \pi^2 A}{T^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3\,a)$$

T, the period, can always be observed; and A can in many cases be easily calculated from a knowledge of the mass and dimensions of the magnet; so $R M$ may be determined.

If the same magnet is allowed to vibrate in succession in two fields of force, whose intensities are R_1 and R_2; and if the periods of vibration are T_1 and T_2,

$$R_1 M = \frac{4 \pi^2 A}{T_1^2},$$

$$R_2 M = \frac{4 \pi^2 A}{T_2^2}.$$

Hence $$R_1 : R_2 = T_2^2 : T_1^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

So by means of a vibrating magnet it is possible to compare the intensities of two fields of force.

270. Measurement of M / R. If this same bar-magnet, whose period has been determined when vibrating freely in a uniform field of intensity, R, is placed at rest at right angles to the field, it itself will produce a field of force which at some distance away is nearly uniform, and at right angles to the existing field, R. So, if a *small* magnetic needle is suspended at a point some distance away in the direction of the length of the bar-magnet in such a manner as to be free to turn around an axis, perpendicular to the field of force and to the bar-magnet, it will be under the influence of two fields of force at right angles to each other. It will therefore place itself at such an angle that the moments due to the two fields are equal but opposite, so that they neutralize each other. Call, for the time being, the intensity of the field due to the bar-magnet, f. The moment due to the field whose intensity is R, when the magnet makes with its direction the angle a, is $R M' \sin a$, if M' is the magnetic moment of the magnetic

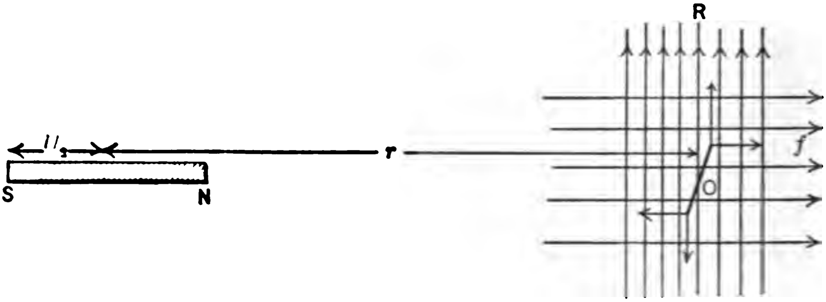


FIG. 197.

needle. But the magnet makes the angle $90° - a$ with the field of intensity f; hence the moment due to it is $fM' \cos a$. These moments are in opposite directions, and must be equal, if the needle is at rest.

Hence $$RM' \sin a = fM' \cos a,$$
or $$\tan a = f/R \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

The value of f may, however, be easily expressed in terms of M, the magnetic moment of the large bar-magnet and r the distance from the centre of this magnet to the centre of the magnetic needle, f is the intensity at the point O, the centre of the needle; i. e. it is the force which would act on a unit north pole if placed there. The large bar-magnet, as placed in the diagram, has a pole of strength $+m$ at a distance $r - \frac{l}{2}$ away from O, and another of strength $-m$ at a distance $r + \frac{l}{2}$. Hence the force on a unit north pole at O is

$$f = \frac{m}{\left(r - \frac{l}{2}\right)^2} - \frac{m}{\left(r + \frac{l}{2}\right)^2} = \frac{2\,m\,l\,r}{\left(r^2 - \frac{l^2}{4}\right)^2}.$$

But if r is very great in comparison with l,

$$f = \frac{2\,m\,l\,r}{r^4} = \frac{2\,M}{r^3} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Consequently, substituting in (5),

$$\tan a = \frac{2M}{Rr^3},$$

or

$$\frac{M}{R} = \frac{r^3 \tan a}{2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (7)$$

Various precautions and modifications for this experiment are explained in laboratory manuals, but it is evident that both r and a can be measured; and so $\frac{M}{R}$ may be determined.

271. Measurement of R or M. By a combination of the two formulæ (3 a) and (7), it is seen that

$$R^2 = \frac{8\pi^2 A}{T^2 r^3 \tan a}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (8)$$

$$M^2 = \frac{2\pi^2 A r^3 \tan a}{T^2}, \quad . \quad . \quad . \quad . \quad . \qquad (9)$$

and so both R and M may be measured.

If R is known for any one field, it has been explained how its value for any other field can be determined by means of a vibrating magnet whose period can be measured.

CHAPTER VI

MAGNETISM OF THE EARTH

272. THAT there is a magnetic field of force around the earth is proved by the fact that a suspended magnet always tends to take a definite position with reference to it. If a magnetic needle is suspended so as to be perfectly free to turn around a vertical and also a horizontal axis, it will not come to rest in a horizontal plane, but will be inclined to it by a certain angle. The direction taken by the needle is, of course, the direction of the line of magnetic force at that point due to the earth. To describe this direction, it is necessary to know how far east or west of the true geographical north and south line the magnet points; and also what angle the magnet makes with the horizontal plane.

Declination. The plane which passes through the centre of the earth and the magnetic needle when it has come to rest is called the "magnetic meridian;" and the angle between this plane and the geographical meridian at the point where the magnet is suspended is called the magnetic "declination" at that point.

It is determined by suspending a magnetic needle so that it is free to turn around a vertical axis, and by then measuring the angle between the direction of the needle and the geographical north and south line. A horizontal line at right angles to the magnetic meridian is said to have the direction "magnetic east and west."

Inclination or Dip. The angle made by the horizontal plane and the needle, which is perfectly free to turn, is called the magnetic "inclination," or "dip."

It is determined by suspending a magnet on a horizontal axis which is placed perpendicular to the magnetic meridian, and then measuring the angle between the magnet and the horizon.

Intensity. The earth's magnetic field has, at each point on the surface, a certain intensity; and the direction of the force is given by the declination and dip. The force at any point may be resolved into two components, — one vertical, the other horizontal. If θ is the dip, R the intensity of the earth's force, H and V the horizontal and vertical components of R, then

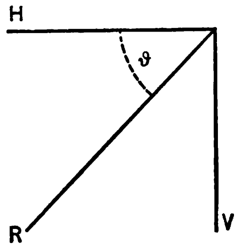


FIG. 198.

$$\left.\begin{aligned} H &= R\cos\theta \\ V &= R\sin\theta \end{aligned}\right\} \quad \ldots\ldots\ldots \quad (1)$$

So that, if θ and H are known, R can be calculated. Similarly,

$$\tan\theta = V/H \quad \ldots\ldots\ldots \quad (1\,a)$$

and, if this ratio V/H can be measured, θ, the dip, is at once determined.

It is perfectly possible to measure H, the horizontal component of the intensity of the earth's field, by the method described in the last chapter (Art. 271). It is simply necessary to perform two experiments: measure the period of a bar-magnet vibrating about a vertical axis; measure the horizontal deflection of a small magnetic needle which is produced by the first magnet when it is placed magnetic east or west of the needle.

Earth's Elements. Variations. The declination, dip, and intensity of the magnetic field at any point on the earth's

surface are known as the magnetic "elements" at that point. They are not constant, but are continually changing. Some of these changes are fairly periodic, while others are not. There is a periodic change each day, each month, and each year; and then there is a secular change of such a kind that the magnetic elements are gradually passing through a cycle which may take several hundred years to complete. Again, there are sudden variations, which are extremely violent, and are said to be due to "magnetic storms."

The explanation of the magnetism of the earth is by no means understood at present; and all that can be done is to study its field of force. There is really no such point as a magnetic pole of the earth; but as ordinarily defined, the magnetic north pole of the earth is that point on the northern hemisphere where a magnetic needle would point vertically down. The magnetic south pole of the earth may be defined in a similar way. It must not be thought, though, that, at any other points on the earth's surface, a magnetic needle points towards a pole of the earth as so defined; a magnet always takes the direction of the line of force, it does not point towards any pole.

CHAPTER VII

MAGNETIC PROPERTIES OF STEADY ELECTRIC CURRENTS

As previously explained (Chap. IV., Art. 253), a steady electric current has an influence on a magnet which may be described by saying that a north pole tends to pass around the current in a "right-handed screw" direction, while a south pole tends to pass in the opposite direction.

273. Magnetic Field around a Current. This law may then be expressed by saying that around any current there is a magnetic field, and that the lines of force are closed curves around the current in a right-handed screw direction. Thus, the relation between direction of current and lines of force is shown in the accompanying figure: they form closed loops encircling each other. In this case that side of the area enclosed by the conductor, which is turned towards the eye, has the properties of a magnetic south pole, because lines of force are entering there; while the opposite side has the properties of a magnetic north pole. For this reason one side of a conducting circuit carrying a current is sometimes called its "south side;" and the other, the "north side." If the area of the circuit is made very small by twisting the wires around each other, the magnetic field due to the current almost completely vanishes.

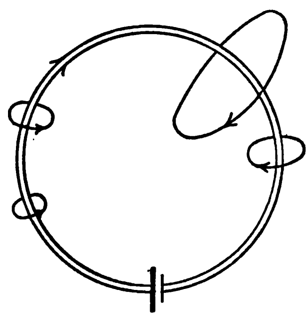


FIG. 199.

Solenoid. A special form of conductor is a helix or spiral; and, if there is a current in it, the magnetic lines of force evidently enter one end, and passing out the other return outside to form closed curves. When this form of a conductor carries a current, it is sometimes called a

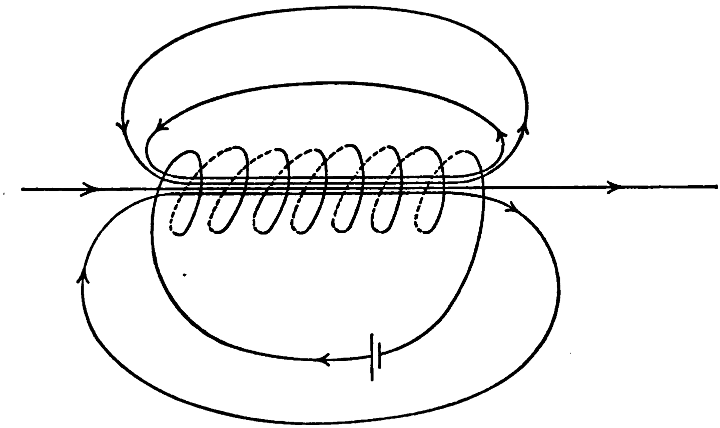


FIG. 200.

"solenoid;" and it has all the external properties of a bar-magnet. So, if a solenoid is suspended free to turn, it will point magnetic north and south. Two solenoids carrying currents act on each other exactly like two bar-magnets.

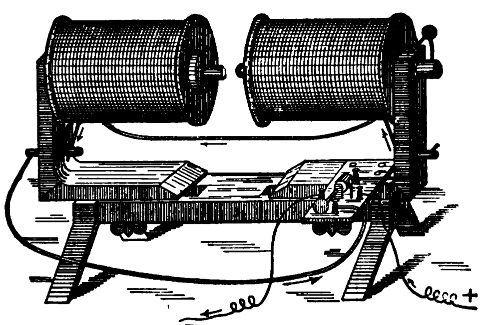


FIG. 201.

Further, it is evident why a bar of iron or steel becomes a magnet when it is placed inside such a solenoid; for the lines of magnetic force come in at one end and leave at the other, thus making it a magnet. If the bar is steel, it remains a magnet; but, if it is iron, it ceases to be one when the current stops. A solenoid with a soft iron bar inside is sometimes called an "electromagnet." A common type of one has its north and south poles facing each other, as shown in Figure 201.

274. Electro-magnetic Force. Since electric currents produce magnetic fields around them, there must be actions between currents, and also between a current and a magnet. (The action of solenoids is an illustration.) The motions will always tend to take place in such a way as to produce a decrease in the potential energy of the magnetic field; and one law has been found which applies to every action, if a magnet may be regarded as a solenoid: the motion always tends of itself to take place in such a way that the field of force passing from the south to the north side through a circuit, carrying a current, increases; or, what is the same thing, the field of force passing from the north to the south side decreases.

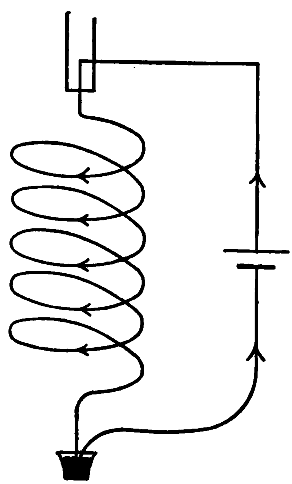


FIG. 202.

Thus, two parallel circuits, carrying currents in the same direction, attract each other; because by coming nearer together more lines of force pass out at the north side of each. This fact is well shown in an experiment which consists in allowing a current to pass through a vertical spiral so hung as to have its lower end just dip into a basin of

mercury. The current in the spiral makes connection with the cell through the mercury. As a consequence of the attraction of the parallel coils which have parallel currents, the spiral contracts and so breaks connection at the mercury surface; the current is thus broken, and the helix falls; it thus makes contact again at the mercury, and then the process repeats itself.

A solenoid, with the current as shown, has its south side at the left, and its north at the right. So, if a bar-

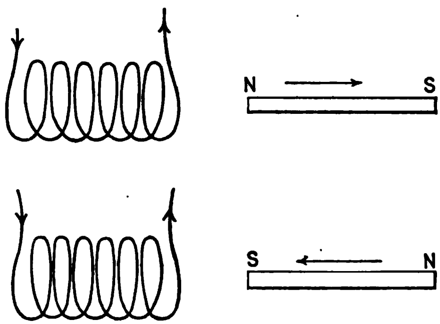


FIG. 203.

magnet is placed at the right of the solenoid with its south pole towards the latter, there will be attraction; because, if the two approach each other, more lines of force come out the north side of both solenoid and magnet. If either current or magnet is reversed, there will be repulsion.

Two parallel circuits, carrying currents in opposite directions, will repel each other, because, when they are close together, one almost neutralizes the magnetic field of the other; and by moving apart the field of force coming out the north side of each will be increased.

275. Ampère's Theory of Magnetism. Ampère proposed an electric theory of magnetism which is exceedingly simple and satisfactory in many respects. Magnetic substances consist of molecules which are themselves magnets; and Ampère's theory is simply that each molecule of a mag-

netic substance has in it an electric current, which probably flows in a fixed channel. If the substance is not a magnet, these molecules are standing at random; but, if a bar of the substance is made a magnet, the process consists in the turning of all the molecular currents until they are more or less parallel. This theory explains at once why a bar of magnetic material becomes a magnet when placed inside a solenoid, because each molecular current tends to turn and place itself parallel to the coils of the helix, so as to have as many lines of force as possible pass through from its south to its north side. This explains, too, why a bar-magnet has the same external action as a solenoid, because, considering the ends of the magnet, lines of force are passing in and coming out through the molecular currents just as if there was a helix wound around the bar, and this helix was carrying a current in the proper direction.

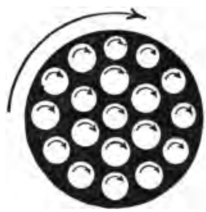


FIG. 204.

276. Electric Motors. The principle of the ordinary electric motor, so common in street cars, elevators, etc., is simply an application of the general law explained above. Coils of wire are arranged so that they can revolve in the field between two magnetic poles. A current is passed from outside through these coils, which then turn so as to have as many lines of force as possible pass out from their north sides. By an automatic device, the current is so regulated and directed that each of the coils is tending to turn; and, as the coils are rigidly fastened to a shaft, this shaft will revolve. By fastening gearing or belting to this shaft, the motor may be made to do work. A particular type of motor will be explained later in Chapter VIII. Article 290.

277. Unit-Current. There is, of course, a connection between the intensity of the electric current and the inten-

sity of the magnetic field produced by it. It may be proved by experiment that, if a current is passed around a circuit of a conductor made in the form of a circle, of radius r, the intensity of the magnetic field at O, the centre of the circle, varies as the intensity of the current, and inversely as the radius. If i is the intensity of the current, and f the intensity of the magnetic field at O, f is proportional to $\frac{i}{r}$; or, writing c for a factor of proportionality, $f = c\frac{i}{r}$. f can be measured as previously explained (see Arts. 271 and 269); so can r, because it is a distance. But the numerical value of i is entirely arbitrary, depending upon the value of c, which is *any* constant. If it is wished, c can be made equal to 2π, in which case

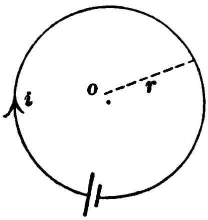


FIG. 205.

$$f = \frac{2\pi i}{r}; \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

and this equation fixes the numerical value of the intensity of that current which, when passed around a conductor bent in the form of a circle of radius r, will produce the magnetic intensity f at the centre. This is equivalent to giving a definition to a current of intensity one, i. e. to a "unit-current;" because, if $f = 1$ and $r = 2\pi$, then $i = 1$. This means that a unit current is such that, if passed around a conductor bent in the form of a circle whose radius is 2π, it will produce a magnetic intensity 1 at the centre. This unit current is called the "electro-magnetic unit;" and the system based on it is called the "electro-magnetic system." A unit quantity of electricity on this system would be the quantity carried by a unit current in one sec. (or by a current of intensity 2, in $\frac{1}{2}$ sec., etc.) It has been found by experiment that there are almost ex-

actly 3×10^{10} electrostatic units of quantity in one electromagnetic unit.

278. Tangent Galvanometer. If a wire is wound so as to form n circular coils, of the same radius r; and, if the helix is so compressed that the coils practically coincide, the magnetic intensity at their centre, produced by a current of intensity i, will be n times as great as that due to a single coil. In this case, then,

$$f = \frac{2\pi n i}{r} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1\,a)$$

The direction of the magnetic force at the centre is, of course, perpendicular to the plane of the coils; so that, if the coils are vertical, and if a small magnetic needle is suspended at their centre, it will tend to point at right angles to the coils. But, if the coils themselves are in the magnetic meridian, the needle at the centre is also acted upon by a moment, due to the earth's field, which tends to place it in the plane of the coil. Consequently, the needle will be deflected, and will come to rest at some angle where the moments due to the fields of the earth and the current balance each other.

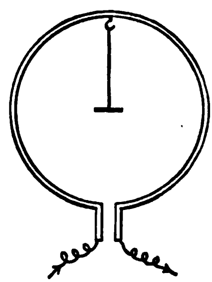


FIG. 266.

Such an apparatus is called a "tangent galvanometer," for reasons which will appear in Formula (2). It consists of a wire wound in several circular coils placed close together, with their planes in the magnetic meridian, and a magnetic needle at the centre of the coils, supported on a vertical pivot. The radius of the coils must be immensely larger than the length of the needle, so that the intensity of the magnetic force due to the current may be the same at the two ends of the needle as at its centre. Let the magnetic moment of the needle be M, and when a current

passes through the coils, let it come to rest at an angle θ with the magnetic meridian. There are two fields of force acting on the needle, whose intensities are H, the horizontal intensity of the earth's field, and $f\left(=\frac{2\pi n i}{r}\right)$ the intensity due to the current. These two fields are at

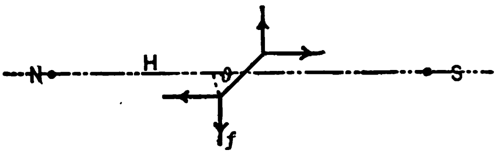


FIG. 207.

right angles to each other; so that the two mechanical moments acting on the magnet are

$$HM \sin\theta \text{ and } fM\cos\theta.$$

These two moments must be equal, since the needle is at rest; hence

$$HM\sin\theta = \frac{2\pi n i}{r} M \cos\theta,$$

or

$$i = \frac{H}{\frac{2\pi n}{r}} \tan\theta \quad . \quad . \quad . \quad . \quad . \quad (2)$$

$\frac{2\pi n}{r}$ is constant for a given instrument, and is called the "galvanometer constant." Writing

$$\frac{2\pi n}{r} = G,$$

the above equation becomes

$$i = \frac{H}{G}\tan\theta \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2\,a)$$

Therefore, if H, G, and θ can be measured, the numerical value of i is determined, and for this reason the instru-

ment is called a "galvanometer." In any case, even if H and G are unknown, the values of the intensities of two currents may be compared by finding the deflections produced by them. Or, if i can be measured by a voltameter (see Art. 246), and G and θ also measured, the value of H may be determined for the place where the galvanometer stands.

This same instrument also furnishes a method for the comparison of the E. M. F.'s of two primary cells. For, by Ohm's law, $i = E / R$; and, if the two cells, having E. M. F.'s E_1 and E_2, are each allowed to produce a current in turn through the *same* resistance,

$$i_1 : i_2 = E_1 : E_2.$$

But the ratio i_1 / i_2 can be measured by placing a galvanometer in the circuit; and thus the ratio E_1 / E_2 may be determined. Other and better methods are described in laboratory manuals.

279. Electro-magnetic System. The electro-magnetic unit-current and unit-quantity of electricity have been already defined; and using them it is possible to define a unit E. M. F. and a unit resistance in this same system. If, as a current of intensity i flows for t seconds through a conductor, the heat-effect in a given portion of that conductor from A to B is H ergs, it has been proved that $H = E\,i\,t$, if E is the E. M. F. between A and B. But the numerical values of H, i, and t are known; so that of E is fixed. This is obviously equivalent to defining a unit E. M. F.

FIG. 208.

This same amount of heat-energy can be expressed in terms of R, the resistance of the conductor between A and B. For, from Formula (11), Chapter IV., $H = i^2\,R\,t$. The numerical values of H, i, and t are known; so that of R is fixed; and this is equivalent to defining a unit resistance.

These electro-magnetic units of E. M. F. and resistance, as deduced from the unit-current, are inconveniently small; so certain multiples of these units are ordinarily used in practical measurements.

280. Practical Units. The practical unit of E. M. F. is called the "volt;" and

$$1 \text{ volt} = 10^8 \text{ electro-magnetic units.}$$

The practical unit of resistance is called the "ohm;" and

$$1 \text{ ohm} = 10^9 \text{ electro-magnetic units.}$$

The practical unit of intensity of current is called the "ampere;" and

$$1 \text{ ampere} = 10^{-1} \text{ electro-magnetic unit.}$$

(Hence an E. M. F. of 1 volt applied at the ends of a conductor of 1 ohm resistance produces a current of 1 ampere.) The practical unit of quantity is called the "coulomb;" and, of course,

$$1 \text{ coulomb} = 10^{-1} \text{ electro-magnetic unit.}$$

Even these units, although of convenient size, are not well adapted for ordinary daily use. For instance, in order to find the E. M. F. between two points of a conductor, it wonld be most inconvenient to measure the heat-effect produced as the current passes; although it might be done in some cases. Similarly, the resistance could not be conveniently measured. But methods are known by which the ratio of two E. M. F.'s or the ratio of two resistances may be most accurately determined. Therefore, with the greatest exactness, the value of a definite standard E. M. F. has been measured in terms of volts; and the value of a definite standard resistance has been measured in terms of ohms. These standards can be compared with the unknown quantities; and so values of the latter may be determined.

The E. M. F. between the poles of a Clark cell, made in a

particular way, is 1.434 volts if the temperature is 15° C. Specifications may be easily obtained for the construction of the cell; and the rate of change of its E. M. F. with the temperature is known; so there is no difficulty in the use of this standard E. M. F.

The resistance of a column of mercury 106.3 cm. long of uniform cross-section, and containing 14.4521 grams is 1 ohm at 0° C. Mercury is easily obtained in a pure condition; and so a column of it in a glass tube of the proper size may serve for a standard resistance.

As stated before, in Table XIV., when an electro-magnetic unit quantity of electricity is carried through a solution of a silver salt, 0.01118 grams of silver are deposited. This is absolutely true only if the solution is one of silver nitrate in water, special precautions being observed. A coulomb will therefore deposit 0.001118 grams of silver; or, what is the same thing, 96,540 coulombs will deposit an amount of any substance equal to its chemical equivalent. By means of a knowledge of the above figures for a coulomb and a silver salt, the intensity of any current may be measured with any kind of voltameter.

281. Energy of a Magnetic Field. Since a magnet has all the properties of an electric current, in discussing the energy of a magnetic field it will be necessary to speak only of currents. There is no magnetic field around a conductor which is not carrying a current; but, as a current is started, a magnetic field is produced, and work is required to do this. This is obvious, because changes are produced in the surrounding ether and matter. Consequently, a definite amount of work is required to *start* a current, to raise it from zero up to its full value; and this energy is undoubtedly present as kinetic energy in the ether and the connected matter. After the current is started, energy is required to keep it going, owing to the work necessary to overcome the resistance of the conductor, this energy being spent in heat-effects in the conductor.

So every current (and magnet) must be considered as associated with a definite amount of kinetic energy in the surrounding medium. The amount of this energy depends, of course, on the extent and intensity of the magnetic field produced by the current, that is, on the form of the conducting circuit and on the intensity of the current.

If there are two currents near each other (or two magnets), there is an extra amount of energy in the field, as is shown by the fact that one current has an action on another. This energy is potential, and depends upon how much work has been required to bring these two currents near each other, if originally they were an infinite distance apart. If work was done by external forces in bringing the currents into position, this energy is added to that of the two currents; while, if the two currents "attract" each other, the energy is subtracted from that of the currents.

CHAPTER VIII

INDUCED CURRENTS

If a conducting circuit, e.g. a wire bent into a closed curve, is placed in a magnetic field, there will in general be a certain field of force passing through the area bounded by the circuit. This is true whether the circuit carries a current itself or not. Faraday discovered that, if in any way this field of force passing through the circuit is changed, there is instantly a current produced in the conducting circuit, or, if there is a current there already, it will be changed. Faraday called these currents which are thus produced by changing the field of magnetic force "induced currents."

282. Properties of Induced Currents. These induced currents last only as long as the field of force through the circuit is being *changed;* and it is proved by experiment that they are always in such a direction as by their own action to oppose or neutralize the change in the field by which they are produced; the amount of the "induced E. M. F.," to which they may be attributed, may be proved to vary directly as the amount of the change in the field of force through the circuit, and inversely as the time taken for the change to be produced. Further, if there are n turns of wire in the circuit, the induced E. M. F. will clearly be n times as great as it would be for a single coil.

Energy is required for the production of these induced currents; and, by considering certain simple illustrations, the sources of the energy will be clearly understood.

SPECIAL CASES

283. 1. A Current and a Magnet. Let the arrangement be as shown; a magnet turned with its south pole towards the south side of the current. There is a definite field of force passing inside the circuit, due to the current and to the magnet. There will be an induced current if this is in any way changed. Let the circuit be kept fixed, and let the magnet be moved towards it. Work is required as long as there is motion; and energy, therefore, leaves the external body which produces the motion, and is available for the induced current which flows while the magnet is being moved.

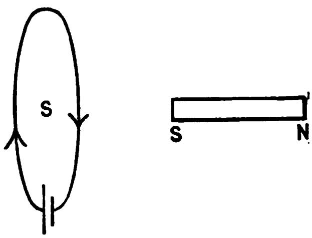


FIG. 209.

The effect on the field of force is to *decrease* the field of force passing through the circuit from the south side to the north; and the induced current will be in such a direction as to *increase* this field; that is, it will be in the same direction as the original current. As soon as the motion of the magnet ceases, the supply of energy ceases; and the induced current dies out rapidly, its energy being spent in heating the conducting circuit.

This induced current presents its south side towards the approaching magnet, and tends to repel it; and the phenomenon may be described by saying that, when the magnet is brought near, the induced current is in such a direction as to repel it, that is to tend to prevent the motion.

Again, suppose that both magnet and circuit are free to move. They will separate, owing to the tendency of the potential energy to become less; but this changes the field of force. As they move apart, the field of force

passing through from the south to the north side of the circuit is *increasing;* so the induced current must be in such a direction as to decrease this field; that is, it is in the opposite direction to the existing current, and the immediate result is a decrease in the actual current.

Work is required to produce the motion of separation of the circuit and magnet; for work is always required to set any body in motion. The only source of energy is the cell which furnishes the current; and, if some of its energy is used up in producing motion, less than usual is available for the current; and consequently the current decreases. As soon as the motion ceases, the current returns to its previous value.

The induced current was in such a direction as to have its north side towards the magnet; that is, it tended to hinder the motion of separation. The phenomenon may, then, be described by saying that the induced current is in such a direction as to tend to oppose the change in position of the circuit and magnet. This law is, in fact, a general one, and applies to all changes in position of currents and magnets.

It is at once evident how these same explanations may be extended to the motions of currents with reference to each other.

284. 2. **Conducting-Circuit and Magnet.** If a conducting-circuit, which has in it no permanent current, is placed near a magnet, as shown, there will be a field of force passing inside the circuit. If this is changed in any way, there will be an induced current. Let the circuit be fixed, and let the magnet be moved nearer. The change in the field of force through the circuit is to *increase* the field coming through the circuit from the side

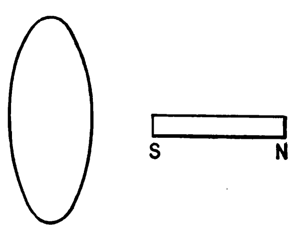


FIG. 210.

away from the magnet to the side towards it; consequently the induced current is in such a direction as to decrease this field; that is, it is such as to make its south side face the approaching magnet. (Here again the induced current is in such a direction as to hinder the change.) Work is done by the force which brings the magnet up, owing to this opposing current; and so the energy is accounted for.

Similarly, if the magnet is moved away, the induced current is in such a direction as to have its north side towards the retreating south pole; and again the motion is hindered. In short, whenever the relative positions of a magnet and a conducting circuit are altered, the induced current is in such a direction as to oppose the change. The energy of the currents is, of course, spent in heating the conductor.

285. A well-known illustration of these laws is "Arago's disc," which is a circular disc of copper (or any conductor) so arranged as to be rapidly revolved in its own plane over (or under) a pivoted magnet. Currents are, of course, induced in the copper disc; and these currents are in such directions as to tend to prevent the motion. Consequently, those portions of the disc which are going away from either pole of the magnet tend to attract it after them so that they would not be separated; portions of the disc approaching the pole tend to repel it. So, as the disc revolves, the magnet tends to follow it, turning in the same direction. The disc, in its turn, becomes hot, as a result of the currents; and the entire energy is, of course, furnished by the power which turns the disc.

An exactly similar phenomenon is observed when a copper disc is pushed between the poles of a powerful electromagnet. In this case, also, the induced currents in the copper are such as to oppose the motion; and work is required to move the disc. If it is moved repeatedly back and forward, its temperature will soon rise.

286. Another illustration of currents in conducting circuits induced by variations in the field of force is afforded by the so-called "earth-inductor." This is a circular framework around which is wound a wire in several turns, and which is so supported as to permit rotation around an axis lying in the plane of the coils, and passing through the centre. Mechanical stops are so arranged as to permit it to turn only 180°, and the ends of the wire are joined to a galvanometer. The apparatus is used in this way: it is so placed that the axis of rotation is vertical, and the plane of the coils is adjusted exactly at right angles to the magnetic meridian; it is then turned rapidly through 180°.

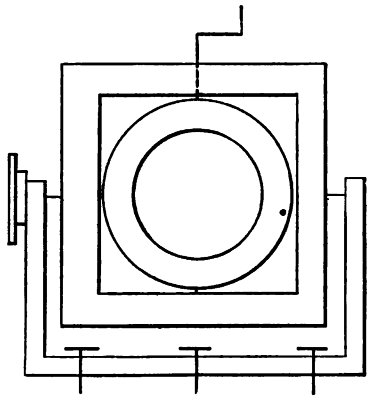


FIG. 211.

When the coils are in their first position, the field of force through them is due to the *horizontal* component of the earth's field; and, when they are turned 90°, there is no force at all through them; but, when turned 90° farther, the horizontal field of the earth again passes through them, but from the opposite side. So, in turning through 180°, the change in the field of force equals twice the field through the coils when they are in their first position. There will, of course, be an induced current; and this will be shown by the galvanometer, which is joined to the coils of wire. It is possible to prove that, if the needle of a galvanometer swings very slowly (1 vibration in about 12 seconds), the extent of the sudden fling of the needle, which is produced by an induced current, is proportional to the induced *quantity*. (Such a galvanometer is called a "ballistic galvanometer," and may

be used to measure any *sudden* flow of electricity. Strictly speaking, the quantity is proportional to the sine of half the angular fling of the needle; but, if the angle is small, the sine of an angle practically equals the angle itself.) The induced E. M. F., and therefore the intensity of the induced current, is, as noted above, proportional to the change in the field of force, and inversely proportional to the time taken. But the induced *quantity* equals the intensity of the current multiplied by the time; and so the quantity is proportional simply to the change in the field of force. Consequently, in the experiment just discussed, the fling of the galvanometer needle is proportional to the horizonal intensity of the earth's field. Let the fling be through the angle a_1, then H equals some constant times a_1, or $H = c\,a_1$.

Now place the coils of wire so that they are horizontal, and rotate them through 180° about the horizontal axis. There will be an induced current, depending upon the amount of the vertical component of the earth's field. Let the fling of the galvanometer be a_2 Then, as above, $V = c\,a_2$,

and, hence, $$\frac{V}{H} = \frac{a_2}{a_1}. \qquad (1)$$

Thus, V/H can be measured; but, by Article 272, Formula (1 a), this ratio is the tangent of the dip; so the dip may be determined.

By means of an earth-inductor and ballistic-galvanometer the intensity of any magnetic field of force, e. g. that between the poles of a dynamo, may be compared with that due to the horizontal component of the earth's field. Therefore, if H is measured absolutely, the intensity of any other magnetic field may be at once determined.

287. 3. A Varying Current in a Conductor. If there is a current in a conductor, there is of course a certain magnetic field around it; so that, if this circuit is broken, the

field tends to disappear. But, as it tends to decrease, there must be an induced current tending to increase it, or rather to keep it constant; and consequently there will be an induced current in the same direction as the original current. This is made evident by the spark seen when a circuit which carries a current is broken. As explained in the last chapter, Article 281, a current is surrounded by a magnetic field which contains energy; the moment the circuit is broken, this energy returns to the conductor which carried the current; and the current really continues for a short time after the E. M. F. of the cell is removed. This induced current is sometimes called the "extra-current on breaking."

Similarly, when a current is being started, the field of force is increasing; and there will be an induced current in such a direction as to oppose the change, that is, opposite to the main current. The effect is that it takes some time to make a current reach its full value. In other words, energy is being given the surrounding medium, and, owing to its inertia, time is necessary for a steady state to be reached. This induced opposition-current is sometimes called the "extra-current on making."

(By way of analogy, compare the fact that, when a railway train starts from rest, some time elapses before it attains its full speed, work being done in producing kinetic energy. After this speed is reached, work is done only in overcoming frictional resistances.)

288. 4. A Varying Current near a Conducting Circuit. This case is in principle exactly the same as cases 1 and 2. In them the field of force inside a conducting circuit was altered by changing the relative *positions* of the circuit and a magnet (or solenoid). This field may be equally well changed by altering the intensity of the current in a solenoid near by. Some of the lines of force due to the solenoid pass through the conducting circuit, in general; and, if the intensity of the current in the solenoid

changes, the field of force changes, and there will be corresponding induced currents in the conducting circuit.

Many instruments have been devised to make use of induced currents; and a few will now be discussed.

289. Transformer, or Induction Coil. An induction coil, or transformer, is an instrument designed to receive a current with a small E. M. F., and to furnish one with a large E. M. F.; or *vice versa*. The principle is exceedingly simple. Two coils of wire are so wound as to have the same magnetic field when a current passes in either. This may be secured by having the two coils wound around the same soft iron ring, even at different portions of it; or by having one coil wound inside the other. One coil consists, in general, of a great number of turns of fine wire; the other, of a few turns of coarse wire.

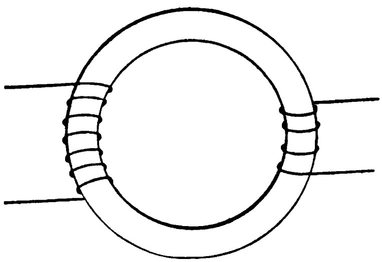


FIG. 212.

When a current is passed through the coarse wire, even if there is a small E. M. F., the intensity will be great because the resistance is small, and so there will be a strong field of magnetic force, especially if there is a core of soft iron inside the coils. This field of force is inside the coil of fine wire, too; and, if now by any means the current in the coarse wire can be either reversed or broken, there will be a great change in the field of force, and a current will be induced in the coil of fine wire. The induced E. M. F. depends upon the rate of change of the field of force, and also upon the number of turns of the wire; because in each turn there will be the same E. M. F. induced, and these are all superimposed so that the E. M. F. at the ends of the coil varies directly as the number of turns of wire. The current in the coarse wire may be

reversed (by an "alternating dynamo") or broken and made again at very frequent intervals; and so the induced E. M. F. in the fine wire will be extremely great. Consequently, when used in this way, a small E. M. F. and large current give rise to a large E. M. F. and a correspondingly small current, because the energy of a current varies as the product of the E. M. F. and the current; and the energy delivered by the apparatus cannot exceed that furnished it.

Conversely, if a very large E. M. F. is applied to the coil of fine wire, it will produce only a small current; but if this is reversed, or broken and made again, at regular intervals, there will be a large current with a small E. M. F. produced in the coil of coarse wire.

To reverse the current in the coil, an alternating dynamo is used; that is, a machine which produces a current first in one direction, then in the other; but simply to make and break an ordinary direct current, any automatic device may be used. The simplest form is shown in Fig. 213 which represents an ordinary "induction coil." The soft iron core of the two coaxial coils extends a short distance beyond the coils; and quite close to its end is placed a piece of soft iron which is fastened on the end of a stiff spring. When this spring is in its natural position, it completes the circuit from the battery through one of the coils of wire; but, when the soft iron core is magnetized by the current, it attracts the soft iron on the end of the spring, and so breaks the circuit. When the circuit is broken, the iron core ceases to be a magnet, the spring flies back to its previous position, contact is made, and the current again flows; and the process is repeated indefinitely. As the field of force through the core is increased and then decreased in succession, it might be expected that the induced current in the other coil would be first in one direction and then in the other. This, in fact, is observed unless special precautions are taken. In an induction-coil,

a condenser is commonly inserted, as shown, in the battery circuit; and this will in general so weaken the E. M. F. on making contact that it produces comparatively little effect. This is done because an induction coil is ordinarily used to produce spark-discharges in one direction between the two

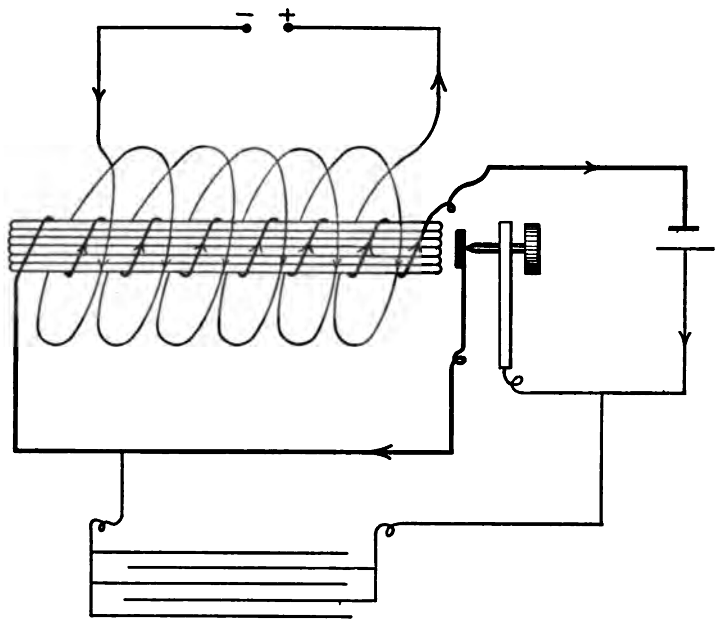


FIG. 213.

terminals of the coil in which the currents are induced; and unless the E. M. F. exceeds a certain limit, sparks cannot pass. It is by means of induction-coils in general, that sparks and discharges through gases are studied. (See Art. 250.)

The iron core of the coils of wire is always made of small iron wires, carefully insulated, placed side by side. If the core was solid, there would be currents induced *around* the iron core itself, as a result of the varying current; and a great deal of energy would be wasted in

heating the iron But, if it is made of wires, insulated from each other, the resistance is so great from one wire to another that a current cannot pass.

290. Dynamos. The simplest case of a so-called dynamo is that of the "Gramme-ring" type. It consists of two parts, — the magnet and the armature. The magnet is one so made that the north and the south poles come opposite

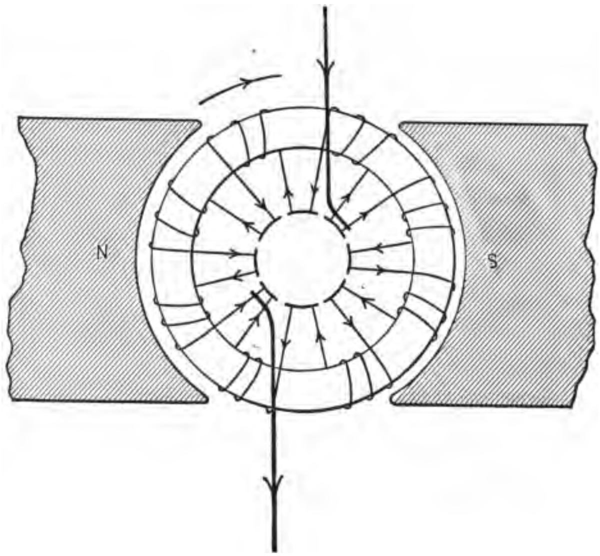


FIG. 214.

each other (as in an ordinary horse-shoe magnet or electro-magnet); and it may be a permanent steel one or an electro-magnet which is magnetized by an electric current. The armature consists of a soft iron ring which is made up of insulated iron wires bent into circles, and around which is wound a continuous copper wire carefully insulated from the iron. The armature is rigidly fastened to a shaft perpendicular to its plane; and the shaft is placed perpendicular to the magnetic field of force between the poles of the magnet. If the shaft is revolved, cur-

rents will, of course, be induced in the coils of wire wound around the ring, because the field of magnetic force through them is constantly changing as the ring revolves. On the shaft of the armature is fastened what is called the "commutator," which consists of metal strips or "bars" along the shaft, each insulated from its neighbors; and resting across these bars are two so-called "brushes," which are metal strips, one on one side of the commutator, the other diametrically opposite, so arranged as to touch opposite bars of the commutator at the same instant. These brushes are held stationary, as the commutator revolves; and they are joined by a conductor, through which a current is desired. Each of the bars of the commutator is joined by a wire to different points of the wire which is wound around the iron ring. Consequently, as the armature is turned by means of the shaft, the brushes are always joined to those portions of the wire around the iron ring which occupy in turn the same positions in the magnetic field.

The lines of force from the north pole of the magnet pass into the iron ring and around through the ring to the side opposite the south pole, where they pass out and cross the air-gap to the pole. They do not pass across the ring, but are divided, as it were, into two sections which crowd through at the top and bottom of the ring. (Compare Fig. 194, Art. 267.) Consequently, as the armature revolves, turns of wire which are horizontal have no field of force through them; but, as they reach the top or bottom and so are placed vertical, there is a strong field through them. Currents are, therefore, induced in these coils; and their directions are easily deduced. If the dynamo is as shown, and the armature is being turned as shown, it is seen that the currents in all the coils on the ascending half of the ring are downward, while those in the other half are also downward. So, if the brushes are connected with those two bars of the commutator which

are joined to the top and bottom coils of the armature, there will be a current produced which will flow from one brush around to the other; then to the top coil, where it will divide into two branches which meet at the bottom coil; and then back to the other brush. This process is perfectly continuous, and a steady current will be produced.

In practice, the brushes are not connected with the coils at top and bottom, but with those a little farther in advance, in the direction of rotation. This is because, as the current is produced in the armature, it produces a magnetic field which so influences the field due to the magnets, that the coil where the current tends to branch is no longer exactly at the top, but is a little in advance.

If, instead of driving the shaft of this armature by means of some external power and so producing a current, a current is sent through the armature from some other dynamo or a battery, the armature will revolve and the shaft can be used to furnish power. This is, of course, the principle of the motor, as already mentioned.

291. Telephone. The ordinary telephone consists of a long steel magnet having a thin coil of wire wound

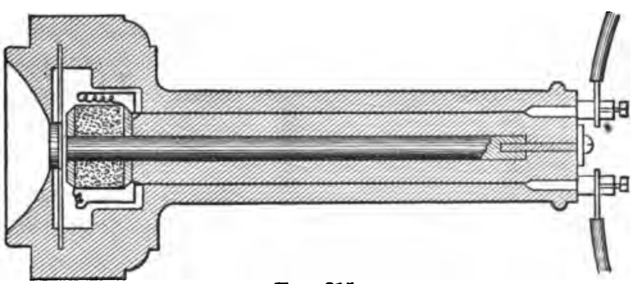


FIG. 215.

around one end, and a soft iron diaphragm held a short distance away from this end. The soft iron plate becomes a magnet under the influence of the steel magnet; and, if it is made to approach the latter, it will alter the field of

force passing through the coil of wire, and there will be an induced current in a certain direction. Similarly, if the plate is moved away, there will be an induced current in the opposite direction. Consequently, if sounds are made near the plate, it will be moved back and forward by the sound-waves; and corresponding induced currents will be produced.

If a current is sent from outside through the coil of wire at the end of the magnet, it will either strengthen or weaken the magnetic field; and the soft iron diaphragm will be either attracted or repelled. If the current was in the opposite direction, the opposite motion would be observed. Consequently, if a current which is being rapidly reversed is sent through the coil, the plate will have corresponding motions back and forward.

If, then, two telephones are joined so that the wire from one coil is connected in series with that from the other, forming a closed circuit, any motion of the plate of one will produce a corresponding motion in that of the other; and any sound made near one will produce corresponding vibrations in the other, which will in turn send out sound-waves.

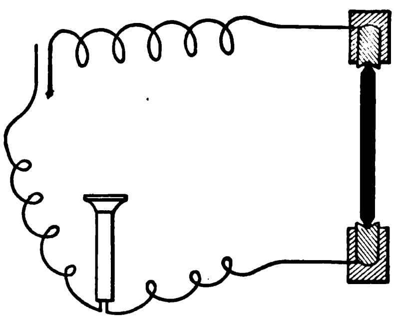


FIG. 216.

292. Microphone. A microphone consists, in principle, of a primary cell joined in series with a "bad-contact;" that is, a junction between two conductors where one rests rather loosely against the other. The resistance of this bad-contact will vary greatly according to the pressure of the two loose conductors; and, if this pressure is changed intermittently by a series of jars or shocks, there will be

corresponding fluctuations in the current flowing from the cell. So, if a telephone is joined in series with the bad-contact, any alteration of the contact will produce a corresponding motion of the telephone diaphragm. In the ordinary microphone used for speaking purposes, the bad-contact is between a plate of metal and a small metallic button back of it; and, if a sound is made near the plate, there are alterations in the contact, and these produce a sound in the telephone which is in series with it.

293. Diamagnetism. There is a theory of diamagnetism, which was proposed by Weber, and which is based upon the properties of induced currents. The fundamental property of a diamagnetic substance is that if placed near a magnet it will be repelled. In Weber's theory it is assumed that the molecules of a diamagnetic substance are perfect conductors, but that they do not have any currents flowing in them, or at least that they are extremely feeble. If such a substance is brought near the north pole of a magnet, there will be an increase in the number of lines entering through the molecules of the substance. Consequently there will be an induced current in each so as to send lines of force *out*; that is, these currents will be in such directions as to have their north sides outward. There will, therefore, be repulsion between the magnet and the diamagnetic substance. If the latter is taken away from the neighborhood of the magnet, currents will be induced in the opposite direction, and the molecules will return to their previous condition.

CHAPTER IX

GENERAL PROPERTIES OF A MAGNETIC FIELD

294. THERE is a great deal of evidence for believing that in a magnetic field there is a rotation or spinning of the minute portions of the ether and matter. This spinning is not motion of large portions of matter, but is just as if a series of molecules were strung on a line, and this line were rapidly rotated around itself as an axis. It is thought that there is some motion like this associated with every magnetic line of force. That rotation of some kind is connected with magnetic fields is shown by two experiments.

295. Rotation of Plane of Polarization. It will be shown in LIGHT that ether-waves can be polarized so that all the transverse vibrations take place in one direction; and such waves are said to be plane polarized. If such a train of ether-waves is sent through a magnetic field, in the direction of the lines of force, it is observed that the direction in which the vibrations are made slowly rotates. This rotation is connected with the direction of the lines of force by the right-handed screw law; and it is most pronounced in very strong fields and in heavy pieces of glass.

296. The "Hall Effect." If a current is passed through a thin sheet of metal, such as foil of some kind, the current spreads out and flows across in certain definite directions. It is always possible to find on one edge of the thin piece of metal a point which has the same potential as a definite point on the other edge; and, in fact, lines can be drawn across the sheet such that all the points on any one

line have the same potential. This can, of course, be done by means of a galvanoscope, which would detect any difference of potential. Suppose that two such points of equal potential on opposite edges of the sheet are joined to a galvanoscope; there will, of course, be no current through it. But it has been found by experiment that if this thin sheet carrying a current is placed in a strong magnetic field perpendicular to the plane of the sheet, e. g. between the poles of an electro-magnet, a current will flow through the galvanoscope. This proves that the two points on the sheet, which had the same potential before, no longer have. In other words, as a result of being placed perpendicular to a magnetic field, the equipotential lines across the sheet are rotated around an axis parallel to the field, so as to lower the potential of the points on one edge of the sheet and raise those on the other.

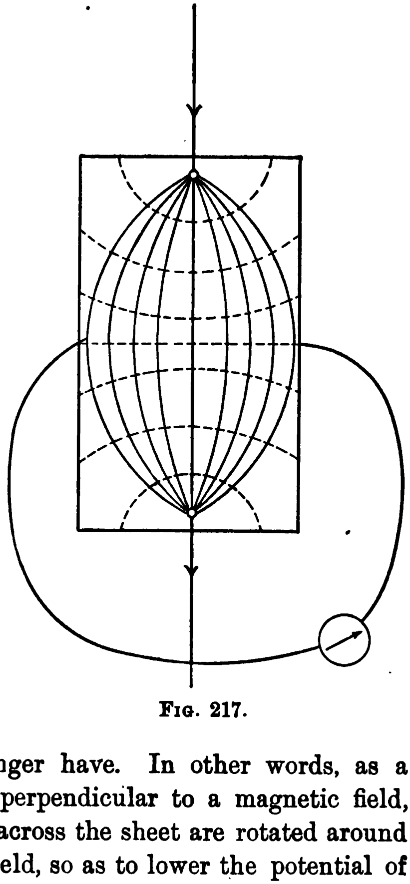


FIG. 217.

BOOK V

LIGHT

BOOK V

LIGHT

INTRODUCTION

297. WHEN the nerves of the retina of the eye are stimulated, the sensation is commonly described by saying that "light" is seen. Light is, then, a pure sensation; and it may be caused in several ways. Ordinarily, it is produced as the result of some action which is taking place some distance away, — such as the combustion of gas in a lamp-flame, or the chemical processes on the sun.

Various colors are distinguished, depending upon the nature of the light-sensation; and variations in the intensity of the sensation are also noticed.

These differences are not confined to the sensations in the eye. Those causes which produce different color sensations also produce different physical effects in many cases, and have different properties. Further, if two sources of light, e. g. a candle and a gas flame, produce sensations of different intensities in the eye, they may be proved to have different physical properties. The source which causes the stronger sensation will produce a greater brightness on any screen which it illuminates, or, if it produces "diffused light" through any semi-transparent substance like paraffin, it will cause a greater illumination than the weaker source. In order, then, that two sources of different strengths may produce equal illuminations in

a block of paraffin, the stronger one must be placed farther off than the weaker. (This statement must be limited somewhat in certain respects, as will appear later.)

A substance is said to be "opaque" if it allows no effect to pass through it; "transparent," if it does. Applied to "light," an opaque body shuts off all illumination; or, if opaque to only certain colors, it prevents the effect passing which can produce those sensations. Similarly a transparent body allows all or only certain light-effects to pass.

It is also known to all that sharp light-shadows are cast, if an opaque obstacle of any kind is interposed between a small source of light, such as a candle, and a screen. This is ordinarily expressed by saying that "light travels in straight lines." Again, light, considered as an effect sent out by some source, is reflected by mirrors and by bodies of all kinds. When light falls upon a rough body, each point of it becomes, as it were, a new source of light; and it is owing to this fact that most natural objects are seen. Almost every one is probably familiar with the fact that, when light, still considered as an effect sent out by a source, falls obliquely upon any transparent body, the direction of its propagation in this body is changed. Illustrations are afforded when one looks into a basin of water obliquely at a body lying on the bottom; also when a lens or prism is interposed between a source of light and a screen.

All these, and many other phenomena associated with light, must be explained, that is, must be shown to be consequences of a simple theory.

298. Ether-Waves. It is now known, as the result of experiments which will be described in Chapter I., that the sensation light is produced when certain waves enter the eye, these waves having been produced in the ether by the actions in the so-called source of light. It will be proved that these waves are transverse, and that waves of differ-

ent periods produce in the eye sensations of different colors. That these waves are in the ether may be considered as demonstrated by the fact that, so far as our experiments enable us to test the point, the presence of ordinary matter is not in the least essential for the propagation of the waves. Further, the waves advance with a finite velocity which can be measured without difficulty, as will be shown in Chapter I., which proves the existence of a medium endowed with inertia. Of course, as might be expected, the presence of ordinary matter in the ether loads it, and so influences the velocity of propagation of the waves, — always making them go more slowly. Waves in the ether may have different lengths depending upon the frequency of the vibration which produces them; but, as will be proved later, the velocity of all waves in the pure ether is the same; while, when there is matter immersed in the ether, the longer waves will be less affected than the shorter ones, and so will have a greater velocity. (There are certain exceptions in strongly absorbing media.)

Of these ether-waves only those which have lengths between rather narrow limits can produce the sensation light; but there are other waves both longer and shorter. Those which can produce light may, of course, be more easily observed and studied than the others, because every normal man is provided by nature with an instrument, the eye, which enables him accurately to observe and measure many of the phenomena of light. Those laws of reflection, refraction, dispersion, interference, polarization, etc., which are deduced by observations on light-waves are also true for longer and shorter waves; the only reason why they are demonstrated for light-waves in preference to the others being the great ease with which these may be observed.

Ether-waves which are too long to produce the sensation light may be studied by various methods described in HEAT, depending upon the fact that all ether-waves are carrying energy. Again, those waves which are too short

to produce the sensation light may be studied by photographic methods, because all ether-waves, and especially the short ones, can produce chemical changes in certain compounds.

299. Properties of Waves. It may be well to give a brief statement of the properties of waves in general; and the student is advised to read the chapters in MECHANICS, SOUND, and HEAT, which explain various kinds of waves. Waves are produced by the vibrations of the source; in ether-waves the matter may be in vibration, and by its action on the ether send out waves in it. If the source is a point, the wave-front will be a sphere; and the intensity of the waves will vary inversely as the square of the distance from the source. (This last fact permits the strengths of two sources of light to be compared and measured, as already explained, by means of a so-called paraffin "photometer.") If the wave-front is a sphere, the waves are called "spherical;" and, if it is a plane, they are called "plane."

The number of vibrations per second of the source is called the "frequency;" and it equals the number of "wave-crests" sent out in one second, or the "wave-number," n. The distance between one "crest" and the next is called the "wave-length," λ; and the distance the waves advance in one second is called the velocity of the wave, v; so that $v = n\lambda$. In a pure medium, e. g. the pure ether, the velocity is independent of everything except the elasticity and inertia.

The primary wave may also be considered as replaced at any instant by secondary waves, as explained in Article 75. This is true of a spherical wave or of a plane wave.

CHAPTER I

THE WAVE-THEORY

By the wave-theory is meant the statement that the phenomena of light (and all those due to other radiations in the ether, heat, chemical, electrical, etc.) are produced by waves in the ether. These waves will be proved later, in Chapter VIII., to be transverse; but for all present purposes it is entirely immaterial whether the waves are transverse or longitudinal. The proof of the existence of these waves and of the intimate connection between them and the sensation light is furnished by any so-called "interference" experiment.

300. Young's Interference Experiment. Three parallel opaque screens are placed some few inches apart; in the first is made a narrow slit or opening with straight parallel edges, S; in the second, there are made two narrow slits parallel to that in the first screen and quite close together, A and B. There is thus a narrow opaque portion between the two slits A and B; and the two screens are so adjusted that when a source of light, such as a candle, is placed in front of the first slit, the candle, the first slit, and the opaque portion between the two slits are in the same straight line. If this is done, the two slits A and B are equally illuminated by light from the slit S, but there will not be uniform illu-

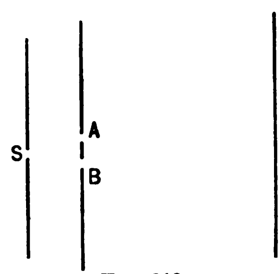


Fig. 218.

mination over the third screen which is receiving light from the two slits A and B. It may be observed that on this third screen there are "bands," parallel to the slits; that is, strips of the screen are illuminated, but in between these strips there is darkness. These bands will be at fairly equal distances apart, if the source of light is an incandescent electric lamp; and, if a piece of colored glass is interposed between the source and the slit S, the bands are at *exactly* equal intervals for quite a wide range. (Instead of having the third screen, an equally good method is simply to receive the light from the two slits A and B in the eye directly.) In any case, it may be necessary to use a microscope in order to see the bands clearly.

The only satisfactory explanation of this experiment is as follows: the ether-waves from the slit S reach the two slits A and B, thus making them two sources of waves, which are identical in all respects, if proper precautions are observed. The two sets of waves which proceed out from the slits A and B illuminate the third screen; but, to reach the same points on the screen, the two trains of waves must go different distances, in general. If there is a point on the screen whose distance from A is r, and whose distance from B is $r \pm \frac{\lambda}{2}$, where λ is the wave-length of the waves emitted by the source, there will be no effect at all at that point, because one train of waves will completely neutralize the effect of the other: the disturbances in a wave at a distance apart of half a wave-length are exactly opposite each other. The central portion of the third screen, that is, the portion in line with the slit S, and the opaque strip between A and B will be illuminated, because the two trains of waves reinforce each other; but on each side of this bright portion there will be a series of points such as described above, where the difference in path of the two trains of waves is $\frac{\lambda}{2}$, and

there is darkness. There is thus a bright central band parallel to the slits, and a dark band on each side. On the farther side of each of these dark bands there will be a bright band made up of points such that their distances from the two slits differ by a whole wave-length λ; for under these conditions the waves strengthen each other. Again, on the farther side of each of these bright bands, there is a dark band, for all the points of which the distances to the two slits differ by $\frac{3\lambda}{2}$, etc. So it is at once evident that by the wave-theory it is possible to completely explain the interference-bands; and on no other theory can it be done.

It is evident, also, that the distance apart of the bands must depend upon the length of the waves; for, the greater λ is, so much the farther must the bright bands be separated, in order that the distances to the two slits may differ by λ, 2λ, etc. It may be proved directly by experiment that, when red light illuminates the slits, the bands are farther apart than when blue light is used, thus proving that red light has a wave-length greater than that of blue light. If all the common colors are examined, it can be proved that they may be arranged in the following order of increasing wave-length: violet, blue, green, yellow, orange, red. If ordinary white light is used, the bands are seen to be overlapping colored bands; and, in fact, the experiment proves that white light is a mixture of waves of all possible wave-lengths (or wave-numbers), which are separated by the interference into their corresponding bands. Any source of light which emits a train of waves of a constant wave-number is said to produce "homogeneous" waves.

(It is obvious that interference-bands are not a phenomenon of *light*, but of waves. Sound-waves can give interference-bands; and so can ether-waves, no matter what their length is: the production and appearance of

the bands is purely a question of the dimensions of the apparatus.)

301. Velocity of Light. These waves which, if of suitable wave-number, can produce the sensation light, are phenomena of the ether, as is proved by many facts. Various stars emit these waves which cause light in our eyes; and there is good evidence for believing that no ordinary matter exists in the space between the stars and the earth; therefore, some medium which is capable of carrying waves must fill this space, and this has been called the ether. That the ether has inertia is proved by the fact that the waves in it travel with a finite velocity. This velocity is very great, being 3×10^{10} cm. per second, or about 186,000 miles per second; but its value may be quite accurately measured. The first observation which led to the determination of the so-called "velocity of light" was the fact that certain discrepancies were noticed between the calculated and the observed times of eclipse of Jupiter's satellites. The planet Jupiter has several satellites, which make regular revolutions around it, just as our moon does around the earth. It seems perfectly easy, then, to calculate when the satellite will disappear behind Jupiter, i. e. be eclipsed; but it was noted that the times of eclipse calculated several months in advance did not agree with the observations. The explanation was most simple: at the end of several months the earth has moved until it is on the opposite side of its orbit from where it was before; and in one of these positions the waves must travel across the diameter of the earth's orbit in order to reach it. This requires some minutes; and the difference between the predicted and the observed times of eclipse may be accurately determined. If the diameter of the orbit is known, and the number of seconds required for the waves to pass across, the velocity of light may be at once calculated.

Another method for the determination of this velocity

was devised by Fizeau, and is known as his method. Light-waves from a slit *S* are allowed to fall upon a glass mirror, *M*, which reflects them at right angles so that they just graze the edge of a large-toothed wheel, *W*. If the wheel is so placed that a tooth intercepts the light, none passes; but, if an opening is in the path, the light-waves pass through, and by means of lenses are directed to a dis-

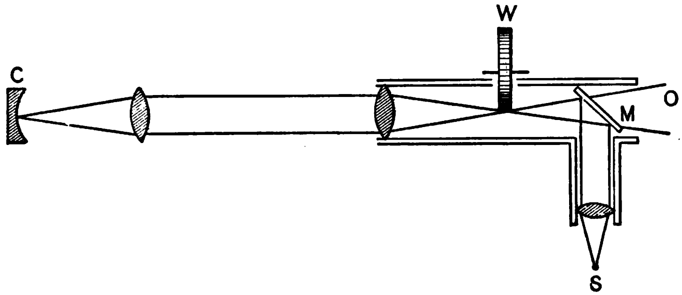


Fig. 219.

tant concave mirror, *C*, which reflects them back over the same path. If the wheel is turned rapidly, it may happen that the waves on their return reach the wheel just as a tooth blocks the way; while, if the rotation is still more rapid, the next opening in the wheel may be in the path, and the return-waves can pass through. They will then fall upon the glass mirror, *M*; and some light will pass through in a straight line, and may be observed at *O*. If the width of a tooth and an opening, the number of revolutions of the wheel per second, and the distance between the wheel and the concave reflecting mirror are known, the velocity of the waves may be at once calculated.

Another and more accurate method is one devised by Foucault, and called by his name. Waves from a source of light at a slit *S* pass through a glass mirror *M*, fall upon a mirror *R*, and are reflected to a distant concave mirror *C*, which reflects them back over their same path; but, just before they reach the slit, they fall upon the

glass mirror *M*, which reflects some of them one side to a point, *O*, where they can be observed. But, if the mirror *R* is revolving about a vertical axis, it will have changed its position slightly during the time the waves took for their passage from *R* to *C*, and back again; and conse-

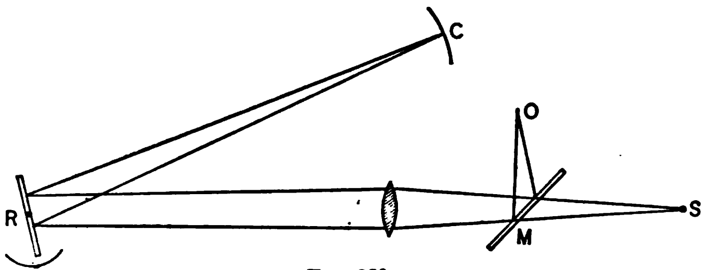


FIG. 220.

quently the return-waves will not be brought to the same point, *O*, to which they would have come if the mirror *R* had been at rest. If the speed of rotation of the revolving mirror *R*, the distance from *R* to *C*, and the shift of the point of light, *O*, are known, the velocity of light may be calculated.

The mean of the best results is, as said above, 3×10^{10} cm. per sec. This is the velocity of ether-waves of all wave-lengths, because, as noted several times (see Arts. 74 and 119), the velocity of waves of any definite kind in a pure medium is independent of the wave-length. For light, this fact is demonstrated by the observation that, when a white star is eclipsed by another, it appears white immediately before and after the eclipse. If red light had a greater velocity than blue, the star would have appeared blue just before it disappeared, and red as soon as it reappeared; but neither this nor the converse is the case.

When the waves enter ether in which there is ordinary matter, the velocity is naturally changed; and Foucault proved by direct experiment that, if water is the matter, the velocity is lessened. This is now known to be true

for all matter, including air; and so a correction must be applied to Fizeau's and Foucault's results in order to determine the velocity in the pure ether. The correction is, however, very small when air is the matter immersed in the ether, being about three parts in ten thousand. As might be expected, waves of different wave-lengths do not travel with the same velocity in ether thus loaded with matter. The general law is, that the less the wave-number, so much the less is the velocity affected. Thus red light is not influenced so much as is blue. In the case of air, the difference for red and blue light is so slight that it cannot be noticed in any ordinary experiments.

302. Rectilinear Propagation. When plane waves are advancing in any direction, they may be considered replaced at any instant, as already said, by secondary spherical waves emitted by each point of the wave-front; and the surface which is tangent to these secondary waves will be the advancing wave-front. The disturbance at any point, *P*, in the path of the wave will then be due to spherical waves sent out by all the points of the primary wave-front which was resolved into secondary waves. It is of importance to learn whether all of these secondary waves influence the disturbance at *P*, or whether some of them neutralize each other. Draw from *P* a perpendicular line upon the primary wave-front, and let it intersect the plane at *O*, which is called the pole of *P*. Call the distance *P O*, *r*. With radii

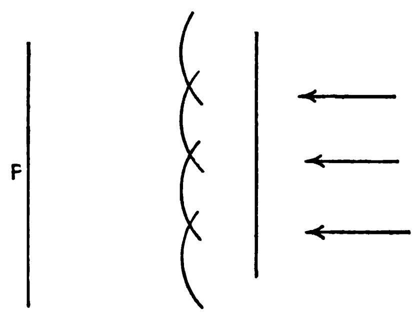


FIG. 221.

$$r + \frac{\lambda}{2},\ r + \lambda,\ r + \frac{3}{2}\lambda,\ r + 2\lambda,\ \text{etc.},$$

where λ is the wave-length of the waves, describe spheres around P as a centre. They will intersect the plane wave-front in circles, as shown, all concentric around the pole O. The plane will thus be divided into zones included between circles; and these are called "Huyghens' zones." It is evident that a spherical wave emitted from a point in any zone will have a distance to pass, in order to reach P, which is exactly $\frac{\lambda}{2}$ shorter than the distance for a definite point in the zone just outside it, and is $\frac{\lambda}{2}$ longer than the distance for a definite point in the zone just inside it. It follows from this that the effect at P due to the secondary waves from any one zone is of the opposite kind from the effects due to the waves coming from the two contiguous zones; for, since the paths differ by $\frac{\lambda}{2}$, the displacements at P will be in opposite directions. The total effect at P due to all the zones, i. e. to the entire series of secondary waves, may, then, be written,

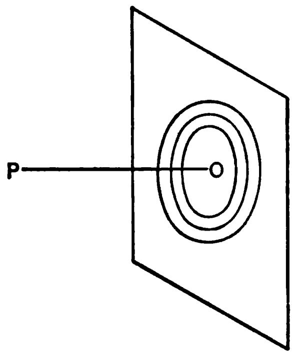


FIG. 222.

$$I = m_1 - m_2 + m_3 - m_4 + m_5 - m_6 + \text{etc.} \quad . \quad . \quad (1)$$

where m_1 is the effect due to the first zone, the central one; m_2, that due to the second, etc. The numerical value of m, the effect due to any zone, depends upon three things: the area of the zone, the distance of the zone from P, the inclination of the waves to the line OP. It may be proved that these first two conditions so neutralize

each other that the effect at P depends only upon the third. The more inclined to the line OP the waves are, so much the less effect do they produce; therefore, the numerical value of any m is greater than that of the succeeding one. The quantity $I = m_1 - m_2 + m_3 - m_4 +$, etc. is, then, a series of decreasing terms, with alternating signs, the differences between consecutive terms being extremely small. The series may be arranged as follows: —

$$I = \tfrac{1}{2} m_1 + \tfrac{1}{2}(m_1 - m_2) - \tfrac{1}{2}(m_2 - m_3) + \tfrac{1}{2}(m_3 - m_4) -, \text{ etc.}$$

But $m_1 - m_2$, $m_2 - m_3$, etc. are so small that

$$\tfrac{1}{2}(m_1 - m_2) - \tfrac{1}{2}(m_2 - m_3) + \tfrac{1}{2}(m_3 - m_4) -, \text{ etc.}$$

is practically zero in comparison with $\frac{1}{2} m_1$. Consequently,

$$I = \tfrac{1}{2} m_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Expressed physically this means that the effect of any one zone is completely neutralized by the action of half of each of the zones on the two sides; so that the entire effect at P is due to the secondary waves emitted by the first central zone around the pole O. The wave-length of ordinary light-waves is about $\frac{1}{45000}$th of an inch; and the area of the first zone must, therefore, be so small as almost to be regarded as a geometrical point.

Corresponding, therefore, to any point in the wave-front, there is a series of points in the path of the waves, the disturbances at which will be due entirely to the secondary waves coming from the first point. If the waves do not pass into a different medium, these points will all lie in a straight line, called a "ray." This phenomenon is sometimes described as the "rectilinear propagation of light." A line drawn perpendicular to the wave-front is called the "wave-normal;" and, if the medium in which the waves are advancing is isotropic, the ray and the wave-normal coincide.

303. Shadows. It is easy, then, to explain why sharp shadows are cast by opaque obstacles, if there is but one

source of light. If a plane wave is advancing from A towards B, and meets there an opaque obstacle, each point of the wave-front O, O', O'', etc., will produce effects at the points P, P', P'', etc. on a screen, C, if $O\,P$, $O'\,P'$, etc., are lines perpendicular to the wave-front. If O is the point on the wave-front just above the obstacle, its corresponding point, P, will be the lowest point on the screen

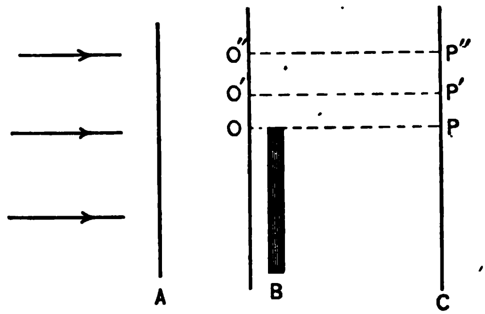


FIG. 223.

to receive waves; and so there should be a sharp shadow below this point. It must be noticed, though, that if Huyghens' zones are drawn around O, as the pole of P, portions of these zones will be below the edge of the obstacle, and so cannot produce any neutralizing effect at P. It would be expected from this that the illumination and general phenomena at the edge of the shadow would not be so simple as at points inside and well outside the shadow. Such is actually the case, as will be explained later. The phenomena at the edge of shadows are included in the name "diffraction" phenomena; but, for general purposes, they may be neglected, and opaque obstacles may be considered as casting sharp shadows. If there are two sources of light, the phenomena are slightly more complicated, because in this case the shadow cast by an opaque obstacle placed in the field of light of one source, S_1, is illuminated in certain regions by the light

from the other source, S_2, and *vice versa.* Other portions of space will receive no light at all, while still others will be illuminated by both sources. This is shown in the diagram, where the black portion, which receives no waves at all, is called the "umbra;" the shaded portions, each of which receives waves from only one source, the "penumbra."

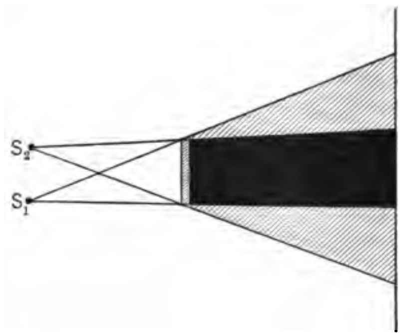


FIG. 224

A still more complicated case is when the source of light is not a point, but of considerable size, like a candle-flame or the sun. If an obstacle is interposed in the path of the waves, there will be, however, as before, umbra and penumbra. A very special illustration is given when the moon passes between the sun and the earth; if the earth is in the umbra, the eclipse is said to be "total;" if in the penumbra, "partial."

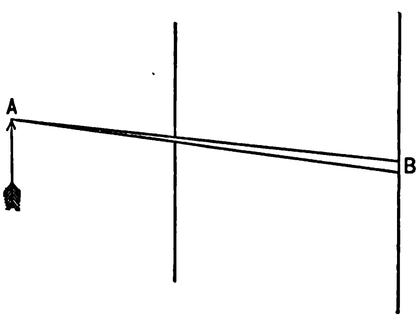


FIG. 225.

304. "Pin-hole Images." Another illustration of the rectilinear propagation of light is afforded by what are known as "pin-hole images." A small opening, e. g. a pin-hole, is made in an opaque screen; and any illuminated object, such as an arrow on which sunlight is falling, is placed in front of the opening. Each point, *A*, of the arrow sends out

waves; but only a small cone of them passes through the pin-hole. A corresponding bright spot is thus formed on a screen at *B*, the spot having the same shape as the pin-hole opening. If the arrow is quite far from the screen which has the opening, and if the screen which receives the light is quite near it, the illuminated spot will be small. Each point of the arrow produces a corresponding bright spot on the screen; and they will all slightly overlap, so that the shape of the pin-hole opening is immaterial. There will, therefore, be on the receiving screen a fairly sharp, inverted image of the arrow. This is the principle of pin-hole photography, where the pin-hole takes the place of a lens; and it also explains the appearance of the images of the sun which are seen under trees whose leaves are thick and close, because in this case the "pin-hole" is some small opening between the leaves.

CHAPTER II.

REFLECTION

As explained in SOUND (Art. 135), whenever waves in one medium reach a surface separating it from another, there will, in general, be reflected waves; and certain important differences were noted between reflection from a medium in which the velocity of propagation is greater than in the original medium, and the similar case when the velocity is less.

305. Mirrors. A mirror is a smooth surface separating two media in which the velocity of the waves is different. "Smoothness" is a purely relative term; but here it is meant to imply that there are no unevennesses which are at all comparable in size with the wave-length of the waves. The commonest forms of mirrors are those whose surfaces are either plane or spherical; and they are called, accordingly, plane or spherical mirrors.

There are several cases of reflection to be considered, depending upon the form of the wave-front and that of the mirror.

306. Plane Waves Incident on a Plane Mirror. Let the incident plane waves meet the plane mirror in a line perpendicular to the plane of the paper; then the section of the plane wave-front may be $OO'O''$, and that of the mirror $O\,Q\,P''$; and their plane is called the "plane of incidence." When the wave-front

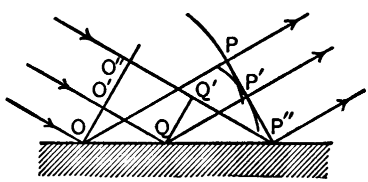


FIG. 226.

has the position $O\,O'\,O''$, it may be considered as replaced by secondary waves; and the wave-front at any future time will be the surface which is tangent to all these secondary waves. The disturbance from O'' proceeds in a straight line perpendicular to the wave-front, as explained in the last chapter (Art. 302), and reaches a point, P'', on the mirror after a certain time. Let the form and position of the wave-front at the end of this time be determined. It must evidently pass through P'' and a line drawn perpendicular to the plane of the paper at that point. Draw from O a line $O\,P$, making the angle $P\,O\,P''$ equal to the angle $O''\,P''\,O$, and from P'' draw a line perpendicular to $O\,P$; then this line may be proved to be a section of the reflected wave-front by the plane of the paper. The triangles $O''\,P''\,O$ and $P\,O\,P''$ are equal; therefore the lines $\overline{O\,P}$ and $\overline{O''\,P''}$ are equal. But while the disturbance has gone from O'' to P'', the secondary spherical waves from O have spread out until the wave-front has an equal radius $O''\,P''$ (or $O\,P$). Therefore, since $P''P$ is perpendicular to $O\,P$, it is tangent to the sphere drawn around O as a centre, and with a radius $O\,P$. Similarly, the disturbance from another point, O', reaches the mirror at Q when the disturbance from O'', reaches Q', where $Q\,Q'$ is a line parallel to the section of the wave-front $O\,O'\,O''$; and, while the disturbance proceeds from Q' to P'', that from Q spreads out in the form of a sphere with a radius equal to $\overline{Q'\,P''}$. If a line $\overline{Q\,P'}$ is drawn perpendicular to $P''\,P$, it is evidently equal in length to $\overline{Q'\,P''}$; hence the plane surface whose section is $P''\,P'\,P$ will be tangent to the secondary wave from Q. Consequently the reflected wave-front is a plane surface which intersects the plane mirror in a line perpendicular to the plane of the paper; and its section with the paper is the line $P''\,P'\,P$.

The "angle of incidence" is defined as the angle between the normal to the incident wave-front and that to the plane mirror, i. e. it is $O''\,O\,P''$. Similarly the "angle

of reflection" is the angle between the normal to the reflected wave-front and that to the mirror, i. e. $P P'' O$. These two angles are evidently equal. Consequently, from these last two paragraphs the laws of ordinary reflection follow at once: —

1. The normals to the incident and reflected wave-fronts and to the plane mirror all lie in one plane, viz. the plane of incidence.

2. The angle of incidence equals the angle of reflection.

These laws, thus deduced on the wave-theory, are verified by direct experiment.

It should be noted that each point in the incident wave-front corresponds to a definite point in the reflected wave-front; so that the effect at the latter point is entirely due to the disturbance at the former. (This can, of course, be proved in detail by drawing Huyghens' zones, as was done in the case of rectilinear propagation.) Thus P corresponds to O, P' to O', P'' to O''. It may be shown, too, that the path $O' Q P'$ is the shortest broken line touching the mirror which can be drawn from O' to P'.

It is hardly necessary to remark that this proof is not a demonstration for a phenomenon of light, but applies to any train of waves falling upon a plane mirror of suitable size. It is true of sound-waves and all others.

307. Rotating Plane Mirror. If plane waves fall normally upon a plane mirror, i. e. if the angle of incidence is zero, waves will be reflected directly back in the same direction. If the mirror is now turned through an angle θ around an axis perpendicular to the plane of incidence, i. e. if the normal to the mirror now makes the angle θ with the normal to the incident waves, the normal to the reflected waves must also make the same angle, θ, with the normal to the mirror, but on the opposite side, because the angle of reflection equals the angle of incidence. Therefore, by turning the mirror through an angle θ, the direction of the reflected waves has been turned through

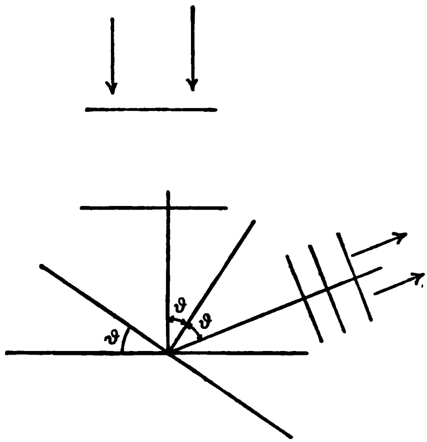


FIG. 227.

an angle 2 θ. This mathematical principle of a rotating mirror is made use of in many instruments, such as the sextant, reflecting mirror-galvanometers, etc.

308. Spherical Waves Incident on a Plane Mirror. If there is a source of waves at a point O, spherical waves will be sent out; and if there is a plane surface near by, the waves will reach it at a point M, where $O\,M$ is perpendicular to the surface. If the surface were not there, the waves would in a certain time reach the position $B\,A\,C$; but, owing to the presence of the surface, when the wave-front reaches M, a secondary spherical wave is sent back, which, in the time taken for the disturbance to reach B and C, has spread to a radius $\overline{M\,A'}$, which equals $\overline{M\,A}$. The reflected wave-front at this instant must include the points B and C, and must be tangent to the sphere of radius $\overline{M\,A'}$ drawn around M. Such a surface, if spheri-

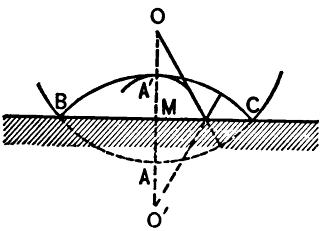


FIG. 228.

cal, must have a centre O' at the same distance perpendicularly below the surface at M as O is above it. It may be proved without difficulty that this spherical surface is really tangent to all the secondary waves emitted by the various points on the mirror around M as the disturbance reaches them; and so it is actually the reflected wave-front.

This is described by saying that, when waves are emitted from a point O and fall upon a plane surface, they *seem* to come after reflection from a point O', where $\overline{O'O}$ is a line perpendicular to the surface and bisected by it. Such a point O' is called a "virtual image" of O; "virtual," because the waves after reflection do not actually come through or from O', but only seem to.

If any illuminated object, e. g. an arrow, is held above a plane mirror, each point of it will have an image in the mirror; and there will be a virtual image of the arrow from which the light will seem to come, if the waves are reflected from the mirror. This image will have the same dimensions as the illuminated object; that is, there is no magnification.

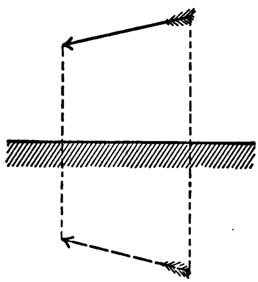


Fig. 229.

Even if the source of waves is to one side of the plane mirror, the same construction and theory applies, because this case is simply a modification of the general solution, made by removing part of the mirror about the point which is perpendicularly below O. When the mirror PQ is to one side, the spherical waves emitted by O pass out in all directions; and some portions of them are intercepted by the surface. The reflected waves are, then, portions of spheres drawn around O' as a centre. The region, though, in which there are any reflected waves is bounded by the cone which can be constructed by drawing straight

lines from O' to each point of the boundary of the mirror PQ. (Of course the phenomena at the edges of the mirror are complicated by diffraction. See Art. 303.) If the area PQ is very small, that portion of the spherical

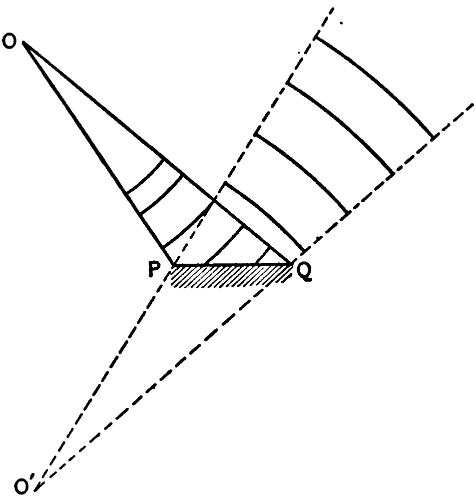


FIG. 230.

waves which is intercepted by it is called a "pencil" or "cone" of waves; and it is evident that the action of the mirror is to change the direction of the cone or pencil, so that it seems to come from the point O' instead of from O.

309. Inclined Mirrors. A special illustration of reflection is afforded by the images formed when a source of light is placed between two plane mirrors which are inclined to each other. The drawing shows the images in the case of a source of light, O, placed between two plane mirrors, *I* and *II*, which make the angle 60° with each other. The images are O', O'', P'', P', P; and they are formed in this way: Some of the waves from O fall upon mirror *I*, and are reflected, leaving the mirror as if they came from O'; but these waves seeming to come from O' fall upon mirror *II*, and are reflected, leaving the mirror as if they came

from P', where $\overline{P'O'}$ is a line perpendicular to mirror *II* and bisected by it; these waves, seeming to come from P', fall upon mirror *I*, and are reflected, leaving the mirror as if they came from P'' where $\overline{P'P''}$ is perpendicular to mirror *I* and bisected by it, etc. Some of the waves from *O* also fall upon the second mirror, and

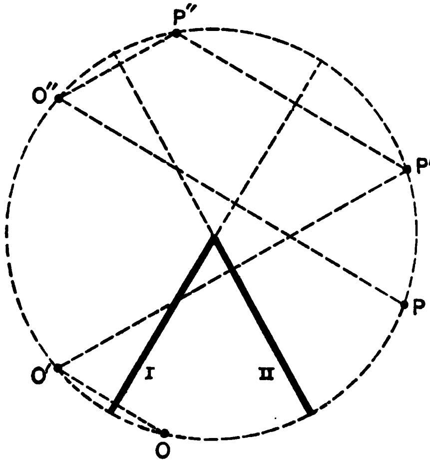


FIG. 231.

have images P, O', P'', etc.; and it may be easily proved by geometry that, if the angle between the mirrors is $\frac{360°}{n}$, where n is any whole number, the images formed in succession in the two mirrors will finally coincide, and that their number is $n - 1$. In the case shown in the figure $n = 6$, and there are 5 images. This is the principle of the kaleidoscope and other toys.

310. Geometrical Measure for Curvature. In considering the reflection of spherical waves from spherical surfaces, a geometrical measure of the curvature of the waves and the surfaces is often useful. A spherical surface is completely defined by its radius, the reciprocal of which is called the "curvature," and it may be easily shown that

there is a simple geometrical method of representing the curvature. Let O be the centre of the spherical surface whose section by the paper is the circle $B\,D\,A$. Draw *any* chord, $A\,B$, and the radius $O\,D$ perpendicular to it. Call $\overline{B\,C} = a$, $\overline{D\,C} = b$; and the angle $B\,O\,D = \theta$.

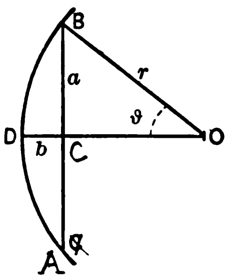


FIG. 232.

Then
$$\overline{B\,C} = a = r \sin \theta$$
$$\overline{C\,O} = r - b = r \cos \theta$$

$\therefore$ squaring and adding,

$$r^2 - 2\,r\,b + b^2 + a^2 = r^2$$

or
$$b^2 + a^2 = 2\,r\,b$$

But, if the curvature is very small, that is, if r is very large and if θ is very small, b^2 may be neglected in comparison with a^2; and so

$$\frac{1}{r} = \frac{2\,b}{a^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

(These conditions are equivalent to assuming that only a small central portion of a spherical surface of large radius is used.) Consequently, if a is constant, $\frac{1}{r}$ is proportional to $\overline{D\,C}$; or, the curvature of any sphere is measured by $\overline{D\,C}$, the "sagitta" corresponding to a chord of fixed length.

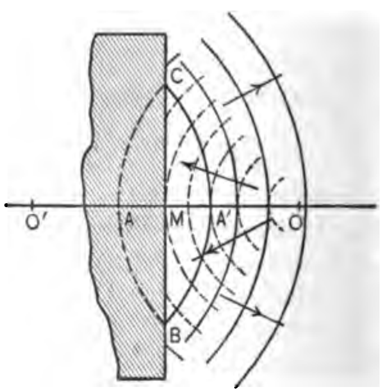


FIG. 233.

Apply this method to the case of spherical waves incident upon a plane surface. The curvature of the incident wave in the position $C\,A\,B$ is measured by $\overline{A\,M}$; that of the

reflected wave by $\overline{A'M}$. But the lines $\overline{AM}$ and $\overline{A'M}$ are of equal length, and in opposite directions. Hence $\overline{A'M} = -\overline{AM}$; and, therefore, a plane surface simply reverses the curvature of an incident train of spherical waves.

311. Spherical Waves Incident on a Spherical Mirror. There are four special cases, according as the waves are diverging or converging, and the surfaces concave or convex. The same demonstration and formulæ, however, apply to them all. The simplest case, perhaps, is that of diverging waves incident upon a concave surface; and it will be considered first.

1. *Diverging Waves, Concave Surface.* Let S be the centre of the concave mirror, and O be the source of di-

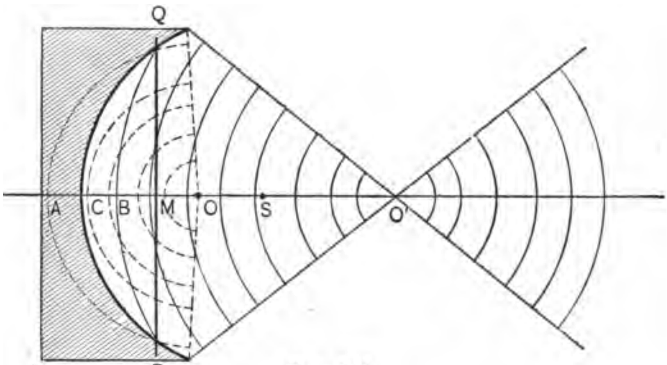


Fig. 234.

verging spherical waves. Draw a straight line through S and O, and let it meet the surface in the point C; also draw *any* chord perpendicular to the line through S and O. This line connecting S and O is called the "axis" of O. If the chord of the sphere meets the axis in the point M, the curvature of the surface is measured by $\overline{CM}$. Draw a sphere around O so as to have the same chord as that of the mirror, and let it intersect the axis in A; then the curvature of the incident waves at a certain instant is

$\overline{AM}$. But, when the disturbance from O reaches C, it does not penetrate into the surface to A, but is reflected back to a point B such that the distance $\overline{CB} = \overline{AC}$; so, when the wave reaches the ends of the chord, P and Q, the reflected wave has the wave-front PBQ, which may be proved to be a spherical surface, if the curvature of the mirror and the portion of it used for reflecting the waves are not too great. So the curvature of the reflected wave is measured by $\overline{BM}$. Writing C, C_i, and C_r for the curvatures of the mirror, the incident wave, and the reflected wave,

$$C = \overline{CM},\ C_i = \overline{AM},\ C_r = \overline{BM}.$$

But
$$\overline{AM} = \overline{AC} + \overline{CM}$$

(remembering that the line $\overline{AM}$ is the line *from A to B*, etc.)

and
$$\overline{CM} = \overline{CB} + \overline{BM},$$

or
$$\overline{BM} = -\overline{CB} + \overline{CM}.$$

Hence, since
$$\overline{AC} = \overline{CB},$$

$$\overline{AM} + \overline{BM} = 2\,\overline{CM},$$

or
$$C_i + C_r = 2\,C \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

In the figure, as shown, the reflected wave is a spherical wave with its centre on the same side as S and O. Let it be at O'. Then the reflected waves will leave the surface and converge towards O'; so that, after contracting there to a point, they will diverge again and spread out in ever-increasing spheres. Consequently, the waves which diverge from O converge, after reflection, at O'; and O and O' are called "conjugate foci." Conversely, of course, waves diverging from O' will converge, after reflection at O.

There are several cases of interest.

312. *a.* If the incident waves are plane, this is equivalent to O being at an infinite distance away on the axis, and $C_i = 0$. Consequently $C_r = 2\,C$; or O' is half-way between S and C at a point F, which is called the "principal

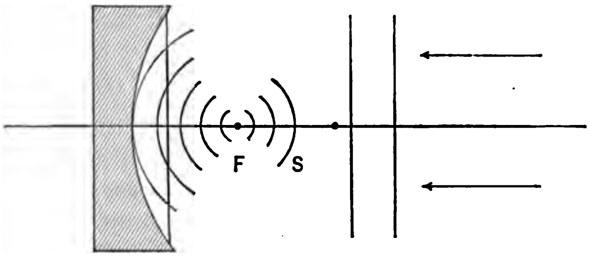


FIG. 285.

focus" for this axis. Conversely, if the source is at F, the reflected waves will be plane waves whose normal is parallel to the direction of the axis of F.

b. If the source is at the centre of the spherical surface, i. e. if O is at S, $C_i = C \therefore C_r = C$; or, waves diverging from S will converge again to the same point after reflection.

313. *c.* If an illuminated object, like an arrow, $O\ P$, is placed in front of a concave mirror, there will be a con-

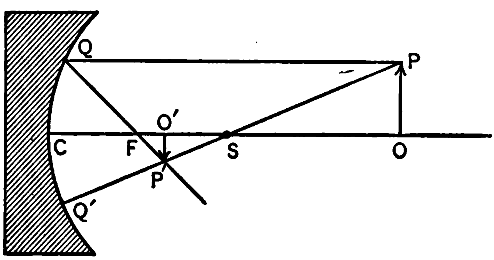


FIG. 236.

jugate focus for each of its points on its corresponding axis; and therefore there will be an image of the entire arrow. The position of this image may be easily found: Let $O\ P$ be perpendicular to $\overline{O\ C}$, the axis of O. Consider parallel waves whose wave-normal is $\overline{O\ C}$ to be advancing towards the mirror; they will reach $O\ P$; and the disturbance from P may be traced, because it is known that

parallel waves are brought to a focus at F, half-way between S and C. Consequently, when P emits spherical waves, that particular disturbance which advances along $\overline{PQ}$ parallel to $\overline{OC}$, the axis of O, must be reflected so as to pass through F along the line $\overline{QF}$. But the spherical waves from P must, after reflection, all converge to some point P' on its axis; that is, P' is the centre to which *all* the disturbances sent out from P must come after reflection. One line of disturbances from P is reflected in the direction $\overline{QF}$; and therefore P' must be at the intersection of the axis $\overline{PQ'}$ with the line $\overline{QF}$, as shown. If the curvature of the mirror is not too great, it may be proved that the image of OP is $O'P'$, perpendicular to the axis of O.

314. *d.* The position and size of the image may also be easily calculated from the above formula, $C_i + C_r = 2\,C$. For C_i is the reciprocal of the distance $\overline{OA}$, i. e. is practically the reciprocal of $\overline{OC}$. Similarly, C_r is practically the reciprocal of $\overline{O'C}$; and C is the reciprocal of $\overline{SC}$. Calling $\overline{OC} = u$, $\overline{O'C} = v$, $\overline{SC} = r$, the above equation becomes

$$\frac{1}{u} + \frac{1}{v} = \frac{2}{r}; \quad . \quad . \quad . \quad . \quad . \quad (2\,a)$$

so that, if r and u are known, v may be calculated.

Further, by geometry, it is evident that

$$\overline{OP} : \overline{O'P'} = \overline{SO} : \overline{SO'} = u - r : r - v,$$

and by equation (2) $\quad u - r : r - v = u : v.$

That is, the *linear* dimensions of object and image are in the ratio of u to v. This ratio of the linear dimensions is sometimes called the linear "magnification."

In the illustration shown, the object causes a *real* image on reflection; because the waves actually converge toward $O'P'$; and, if a screen is placed there, the image will appear on it.

315. *e.* In certain cases the image may be virtual. For, since $C_i + C_r = 2\,C$, if $C_i > 2\,C$, C_r is negative. This

means that, if the object is placed nearer the mirror than the principal focus, the image must be back of the mirror, and so must be virtual. A figure is given to show this, in which an object $\overline{O\,P}$ is placed between F and C; and the virtual image is at $\overline{O'\,P'}$. Hence waves diverging from O

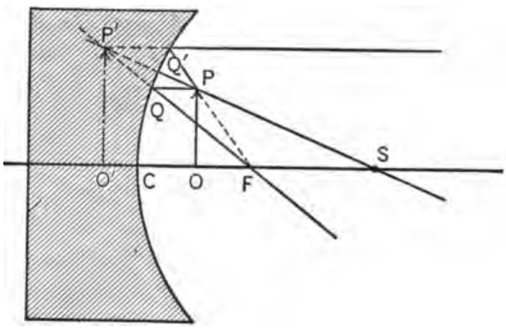


FIG. 237.

fall upon the mirror and are reflected, seeming to come from O'; that is, they diverge after reflection from the virtual image O'.

2. *Converging Waves, Concave Surface.* This is evidently the converse of the case just considered. For, if waves are converging toward O', they will be reflected so as to converge toward O. (This is true whether O' is a real or a virtual image in the previous case.)

It should be remembered that, if waves are converging to a point which is on the opposite side of the mirror from its centre of curvature S, the numerical value of C_i must be given a sign opposite to that of C.

316. 3. *Diverging Waves, Convex Surface.* Let O be the source of the waves, and S the centre of the surface. Draw the axis $O\,S$, and proceed exactly as in Case 1.

$$C = \overline{C\,M}, \quad C_i = \overline{A\,M}, \quad C_r = \overline{B\,M},$$

and

$$\overline{A\,C} = \overline{C\,B}.$$

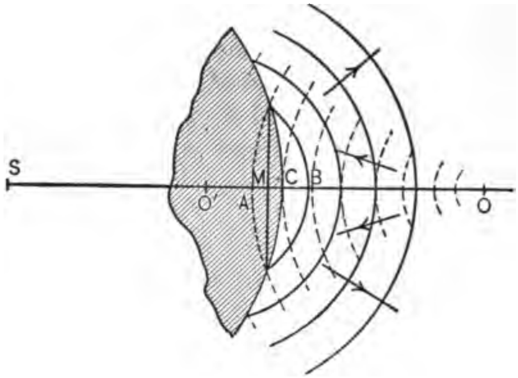


FIG. 238.

But
$$\overline{AM} = \overline{AC} + \overline{CM}$$
$$\overline{BM} = -\overline{CB} + \overline{CM}.$$

Hence, as before,
$$C_i + C_r = 2\,C,$$
or, writing $\quad \overline{SC} = r, \quad \overline{OC} = u, \quad \overline{O'C} = v,$
$$\frac{1}{u} + \frac{1}{v} = \frac{2}{r}.$$

The only precaution is to give the numerical values of C_i and C_r and of u and v proper signs, depending upon whether O and O' are on the same side as S or the opposite.

A special case is when $C_i < 2\,C$. Then C_r is positive, and the image O' is on the same side of the surface as is S. Consequently, waves diverging from O will, after reflection, seem to diverge from a virtual image O'.

4. *Converging Waves, Convex Surface.* This is evidently the converse of the special case discussed in the last section. For, if waves converge toward O', they will, after reflection, converge to O.

317. Spherical Aberration. The theory of spherical mirrors as given in the preceding section applies, as carefully noted, only to small mirrors of small curvatures, or, what

is the same thing, to small portions immediately around the axis, if the mirror is large. If other portions of a mirror than that close to the axis are used, the phenomena are not so simple. Thus, if spherical waves from O fall upon a large concave mirror, those waves which are reflected from the surface near C, the extremity of the axis of O, are brought to a focus at the point O' on the axis.

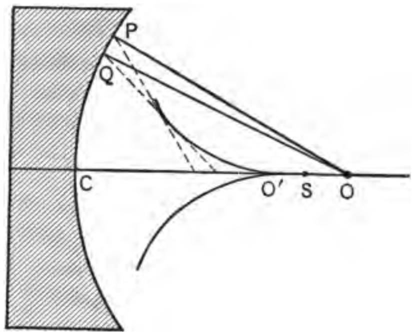


FIG. 239.

But waves which fall obliquely on the mirror, as at P and Q, are not reflected to O', but illuminate a surface whose section by the paper is a pointed curve as shown. This surface, then, is rather conical, with its vertex at O'; and it is called the "caustic" surface. Illustrations are offered when light is reflected by the sides of a tumbler or teacup, and are doubtless most familiar. In general, then, unless the waves fall perpendicularly upon a spherical surface, the reflected waves are not brought to a focus at a point, but illuminate a surface; the entire phenomenon being called "spherical aberration."

CHAPTER III

REFRACTION

318. As explained in the discussion in SOUND (Art. 137), when waves in one medium reach the bounding surface which separates it from another, there are always waves produced in this second medium. So, if ether-waves reach a surface separating the pure ether from a transparent substance, such as water or glass, some waves will enter the ether in which the matter is immersed. The same is true, of course, if the waves reach a surface which separates two transparent substances. The waves are said to be "refracted" when they pass from the pure ether into ether in which matter is immersed, or from one transparent substance into another. The reason for this name will be explained later. It should be remembered that it has been proved (Art. 301) that the velocity of ether-waves is decreased when they pass from pure ether into ether loaded with matter. The wave-number of a train of waves is not changed by refraction, because the same number of waves must go away from the separating surface as come up to it. The velocity $v = n\lambda$; and, since v changes but n does not, λ, the wave-length, must change on refraction.

Several cases of refracted waves will be studied in detail.

319. Plane Waves Refracted at a Plane Surface. Let plane waves of a constant wave-number be incident upon a plane surface separating the pure ether from ether containing matter immersed in it, e. g. water or glass. Let the section of the incident wave-front at any instant be $O\ O'\ O''$; and

let the section of the surface be $O\,Q\,P''$, the line of intersection of the two surfaces being perpendicular to the plane of the paper. Draw the wave-normal $O''\,P''$. At the end of a certain time, t seconds, the disturbance from O'' will reach the surface at P'', where $\overline{O''\,P''} = vt$ if v is the velocity of the waves in the ether. In this same time the disturbance from O has spread out in the lower medium into

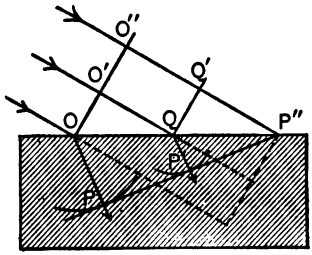


FIG. 240.

a sphere of radius $\overline{O\,P}$ where $\overline{O\,P} = v_1\, t$, v_1 being the velocity of these particular waves in the ether loaded with the matter. The wave-front in this lower medium at this time must include a line through P'' perpendicular to the plane of the paper; and it must also be tangent to the sphere of radius $\overline{O\,P}$ around O as a centre. Draw from P'' a straight line tangent to the circle, which is the section by the paper of the sphere around O. This line is the trace on the paper of the wave-front; for, consider the disturbance from any point O' of the incident wave, which is advancing in the direction $O'\,Q$. It reaches Q when the disturbance from O'' reaches Q', if the line $\overline{Q\,Q'}$ is drawn parallel to $O\,O'\,O''$; and while this last disturbance advances to P'', a wave will spread out in the lower medium around Q of a radius $\overline{Q\,P'}$ where

$$\overline{Q\,P'} : \overline{O\,P} = \overline{Q'\,P''} : \overline{O''\,P''}.$$

But, if a sphere is drawn around Q tangent to the line $O''\,P$, it will have a radius, r, which, by similar triangles, may be expressed

$$r : \overline{O\,P} = \overline{P''\,Q} : \overline{P''\,O} = \overline{Q'\,P''} : \overline{O''\,P''}.$$

Hence $r = \overline{Q\,P'}$; and so the spherical waves sent out from Q will be tangent to the line $\overline{PP''}$, when the disturbance

from O' reaches P''. Therefore, as said, the wave-front in the lower medium is a plane whose section by the paper is the line $P P' P''$; and it is evident that the direction of the wave-front has been changed on entering the lower medium. The waves are, therefore, said to be "refracted" or bent. The angle between the normal to the surface and that to the refracted wave, i. e. $O P'' P$, is called the "angle of refraction," corresponding to the angle of incidence $O' O P''$.

As noted above,

$$\overline{O' P''} = v t \quad \text{and} \quad \overline{O P} = v_1 t.$$

Calling the angle $O' O P'' = i$, and $P P'' O = r$,

$$\overline{O' P''} = \overline{O P''} \sin i, \quad \overline{O P} = \overline{O P''} \sin r.$$

Hence

$$\frac{\sin i}{\sin r} = \frac{v}{v_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

This ratio, $\frac{v}{v_1}$, of the velocity of waves in the pure ether to that of certain waves in ether in which there is a definite kind of matter, is called the "index of refraction" of that matter for the given waves. It is a constant if the wave-theory is true. Consequently, the ordinary laws of refraction follow at once:—

1. The normals to the incident and refracted wave-fronts and to the plane surface all lie in one plane, viz. the plane of incidence.

2. The ratio of the sine of the angle of incidence to the sine of the angle of refraction is a constant for a given form of matter and waves of a definite wave-number; it is entirely independent of the angle of incidence itself.

These laws have been perfectly verified by experiment.

It has been proved by direct experiment that, in all cases when the waves pass from the pure ether into ether contained in transparent matter, $i > r$; that is, $v > v_1$. Therefore, the velocity of ether-waves is always diminished by passing into such matter.

It has been proved further by direct experiment that waves of different wave-numbers are refracted differently, and that blue light is refracted more than green, and green more than red, when the waves pass from pure ether into any ordinary transparent medium. But, as proved in Chapter I., Article 300, red light is characterized by having a longer wave-length (therefore less wave-number) than green; and green has a longer wave-length than blue. Hence, in passing into ether which is loaded with matter, the longer waves are not so much affected as the shorter ones, and are therefore faster. (In other words, waves of small wave-number are less influenced than those of large.) See Article 344.

It follows from the above demonstration that, if the upper medium is not pure ether, but has matter immersed in it, so that the velocity of these particular waves in it is v_2,

$$\frac{\sin i_2}{\sin i_1} = \frac{v_2}{v_1},$$

if i_1 is the angle between the wave-normal and the normal to the surface in the medium where the velocity is v_1; and i_2 is the corresponding angle in the second medium. If μ_1 is the index of refraction for the lower medium, and μ_2 for the upper, i. e. if

$$\mu_1 = \frac{v}{v_1}, \quad \mu_2 = \frac{v}{v_2},$$

then
$$\frac{v_2}{v_1} = \frac{\mu_1}{\mu_2},$$

and
$$\frac{\sin i_2}{\sin i_1} = \frac{\mu_1}{\mu_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

or
$$\mu_1 \sin i_1 = \mu_2 \sin i_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2\,a)$$

Some values of μ for different substances and different waves are given in the following table. Methods of exact measurement will be discussed later. (It should be stated

that μ varies with the temperature, and also with the pressure, in the case of gases; but the variations are slight.)

TABLE XVI

Indices of Refraction

Substance.	Wave-length. cm.	Index.	Temperature. $p = 76$ cm.
Air	0.0000589	1.0002922	0° C
"	0.0000485	1.0002943	0°
"	0.0000434	1.0002962	0°
Helium	0.0000589	1.000043	
Hydrogen	0.0000589	1.000140	0°
Nitrogen	0.0000589	1.000297	0°
Oxygen	0.0000589	1.000272	0°
Alcohol	0.0000589	1.360	15°
Chloroform	0.0000589	1.449	15°
Carbon Bisulphide .	0.0000589	1.624	25°
" "	0.0000485	1.648	25°
Water	0.0000589	1.334	16°
"	0.0000485	1.338	16°
"	0.0000434	1.341	16°
Rock Salt	0.0000589	1.5441	24°
" "	0.0000485	1.5531	24°
" "	0.0000434	1.5607	24°
Flint Glass	0.0000589	1.651	
" "	0.0000485	1.665	
" "	0.0000434	1.677	
Crown Glass . . .	0.0000589	1.517	
" " . . .	0.0000485	1.524	
" " . . .	0.0000434	1.529	

Several special cases of the refraction of plane waves at a plane surface are of interest.

320. *a.* **Total Reflection.** The formula for the angles of incidence and refraction is

$$\mu_1 \sin i_1 = \mu_2 \sin i_2,$$

where $$\mu_1 / \mu_2 = v_2 / v_1.$$

Assume that v_2 is greater than v_1, and that the waves are incident from the medium in which the velocity is less. i_2 is, then, greater than i_1; and, if $P_1 O$ and $O P_2$ are the wave-normals in the two media, i_2 and i_1 will be as shown. i_2 cannot be greater than $90°$; and so the greatest value which i_1 can have if a refracted wave is still produced is that found by substituting for i_2 $90°$. Hence this maximum value of i_1, which may be called a, is given by the equation

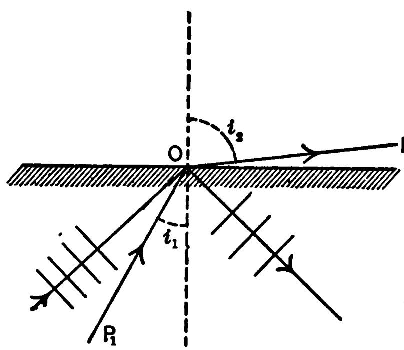


FIG. 241.

$$\mu_1 \sin a = \mu_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

(for $\sin 90° = 1$). If i_1 is increased beyond a, no waves can be refracted; and they will all be reflected, as shown. This angle a, which marks the limiting value of i_1 before total reflection follows, is called the "critical" angle. It is observed, of course, only when waves are passing from one medium into another in which the velocity is greater, e. g. from water into air.

As already noted, μ is different for waves of different wave-numbers; and so it would be expected that different waves would have different critical angles, which is an observed fact. This determination of the critical angle may be made quite exactly; and so this is one method of measuring indices of refraction, for

$$\frac{\mu_2}{\mu_1} = \sin a \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3\,a)$$

321. *b.* Refraction through a Plate with Plane Parallel Faces. When plane waves fall upon a plane surface, the connection between the angles of incidence and refraction is given by

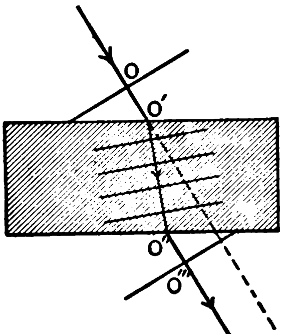


FIG. 242.

$$\mu_1 \sin i_1 = \mu_2 \sin i_2,$$

if i_1 is the angle between the wave-normal and normal to the surface in the first medium, and i_2 the corresponding angle in the second medium. If the second medium is bounded by two parallel surfaces, the refracted waves will fall upon its second surface at an angle i_2. Therefore, calling i the angle of refraction of these waves when they emerge again into the first medium on the other side of the plate,

$$\mu_2 \sin i_2 = \mu_1 \sin i.$$

Hence $i = i_1$; and it follows that the waves emerge parallel to the direction which they had before they entered the plate. A plate with plane parallel sides does not, then, turn the direction of plane waves. But, if the path of the disturbance due to any point in the incident waves is traced, it will be found to be shifted one side parallel to itself. Thus, the disturbance from O travels to O' to O'' to O'''; but the line $O\,O'$, although parallel to the line $O''\,O'''$, is not coincident with it, as is seen from the figure. If $i_1 = o$, however, the two lines do coincide. It follows from this that, if the incidence upon the plate is normal, there is no change in the path of the waves; but, if the incidence is oblique, the emerging waves, although parallel to the incident ones, are, as it were, shifted one side.

322. *c.* **Refraction through a Prism.** A prism is a piece of transparent substance, two portions of whose bounding surface are planes, inclined to each other. If these planes actually intersect in a line, it is known as the "edge" of the prism, and in any case these two plane surfaces are called the "faces" of the prism; and the angle which they make with each other is called the "angle of the prism." If plane waves fall upon one of the faces of a prism so that the plane of incidence is perpendicular to the edge of the prism, the refracted and emerging waves will be as shown in the figure; and it is obvious that the effect of the prism is to change the direction of the waves. But, since the refraction is different for waves of different wave-numbers, it would be expected that a prism would deviate various waves differently, — a fact which is actually observed. (It ought to be stated that it is by means of a prism that this fact in regard to the refraction being different for different waves is most easily proved.)

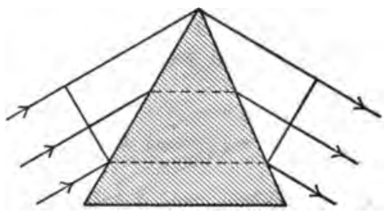


FIG. 243.

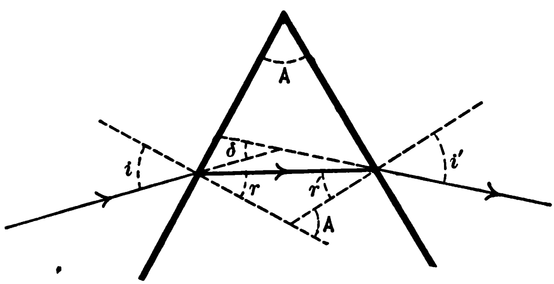


FIG. 244.

The "deviation" of a given prism for a given train of plane waves is the change in direction produced by it in

the waves. Its value may be easily calculated in terms of measurable quantities. Let the directions of the wave-normals and the normals to the faces of the prism be as shown in the figure. Call the angle of incidence i; the angle of refraction r; the angle of incidence on the second face of the prism r'; the angle of emergence i'. If the incident wave-normal is prolonged until it meets the wave-normal of the emerging waves prolonged backward, as shown, δ is the deviation. Call the angle of the prism A. Then, from ordinary geometry, it follows that

$$\left.\begin{array}{l} A = r + r' \\ \delta = (i - r) + (i' - r') = i + i' - A \end{array}\right\} . \quad . \quad (4)$$

So it is evident that δ depends upon the angle of incidence, the angle of the prism, and the index of refraction (because i' depends upon μ).

It may be proved theoretically, as well as by experiment, that for any definite train of waves there is a certain minimum value of the deviation, smaller than which it cannot become for a given prism; and that, further, this minimum deviation occurs when the incident and emerging waves are symmetrical with reference to the prism, i. e. when $i = i'$. But $\dfrac{\sin i}{\sin r} = \mu$ and $\dfrac{\sin i'}{\sin r'} = \mu$; hence, if $i = i'$, $r = r'$. If, then, D is the angle of minimum deviation and i and r the corresponding values of those angles,

$$A = 2\,r,$$
$$D = 2\,i - A.$$

This gives at once a method for the determination of μ; for

$$\mu = \frac{\sin i}{\sin r} = \frac{\sin \dfrac{A + D}{2}}{\sin \dfrac{A}{2}}, \quad . \quad . \quad . \quad . \quad (5)$$

and A and D can both be easily measured.

323. Spherical Waves Refracted at a Plane Surface. The character of the refracted waves may be easily determined for this case, if the definition and measure of curvature given in Article 310 is remembered. As the simplest illustration, consider spherical waves diverging from a point O in one medium; and draw a line OM perpendicular to the plane surface separating the two media. If

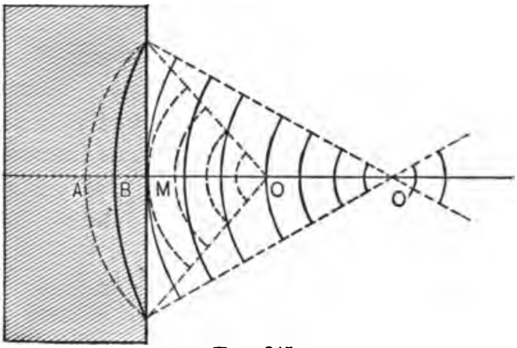


FIG. 245.

the surface was not there, the curvature of the incident waves at a certain instant when the radius was $\overline{OA}$ would be $\overline{AM}$; but, owing to the presence of the second medium, the disturbance which would have gone from M to A really goes from M to B (say), so that the section of the refracted wave is a circle passing through B, and the ends of the chord in which the circle of radius $\overline{OA}$ around O cuts the plane surface. The centre of the refracted waves, i. e. the image of O, is a point O' on the line OM; and the curvature of the refracted wave at this instant is $\overline{BM}$ where $\overline{BM}$: $\overline{AM} = v_2 : v_1$, the ratio of the velocities in the two media.

Hence, writing C_i and C_r for the curvature of the incident and refracted waves,

$$C_r / C_i = v_2 / v_1 = \mu_1 / \mu_2,$$

or

$$C_r = \frac{\mu_1}{\mu_2} C_i \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Therefore the refracted waves diverge as if from a centre O' where $\overline{O'B} = 1 / C_r$. (Conversely, waves in the second medium converging toward a point O' in the first will converge, after refraction, at the point O.) Expressed in terms of distances, $\dfrac{\overline{OA}}{\overline{O'B}} = \dfrac{\mu_1}{\mu_2}$; or, very approximately,

$$\frac{\overline{OM}}{\overline{O'M}} = \frac{\mu_1}{\mu_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

324. A special application of this formula furnishes a method for the determination of μ_1 / μ_2. If there is a source of light at O, a point on the under surface of a plate of a substance which has parallel sides, and whose thickness is d, the waves, diverging from O and refracted out from the upper surface into the surrounding medium, will seem to diverge from another point, O'. (The drawing is made so as to apply to a plate of glass surrounded by air, or to a plate of any transparent substance surrounded by a medium in which the velocity is greater. $\dfrac{\overline{OM}}{\overline{O'M}} = \dfrac{\mu_1}{\mu_2}$, where μ_1 is the index of refraction for the plate, and μ_2 for the surrounding medium. $\overline{OM}$ is the thickness of the plate, d; and, as $\overline{OO'}$ can be easily measured, $\overline{O'M}$ may be calculated, and thus μ_1 / μ_2 determined. The simplest method to measure $\overline{OO'}$ is to focus a vertical microscope on the point O when the plate is not in place, then to interpose the plate horizontally, and measure the distance through which it is necessary to move the microscope in order again to see the object, which is really at O but which seems to be at O'.

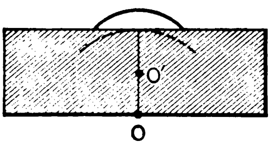


FIG. 246.

If the plate has a refractive index greater than that of the surrounding medium, i. e. if $\mu_1 > \mu_2$, $\overline{OM} > \overline{O'M}$, or the point O' is raised up nearer the upper surface of the

plate. This explains the fact so often observed, that an object at the bottom of a vessel of water never seems to be so far below the surface as the known depth of the water.

325. Another application is to spherical waves falling upon a prism. Let the spherical waves diverge from O; after entering the prism they will seem to diverge from a point O', where the line $\overline{O\,O'}$ is perpendicular to the first

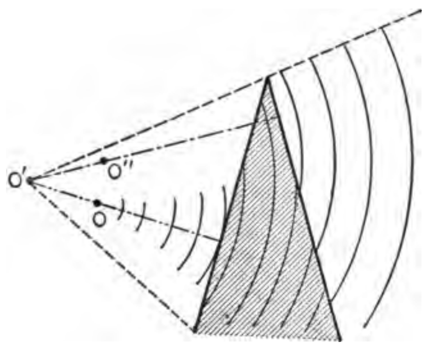


FIG. 247.

face of the prism. On emerging from the prism the waves diverging from O' will seem to diverge from O'' where the line $O'\,O''$ is perpendicular to the second face of the prism. Thus the waves which originally diverged from O leave the prism as if they diverged from O'', so that O'' is a virtual image of O. Since the refraction varies for waves of different wave-numbers, i. e. for waves of different colors if they are light-waves, each train of waves of a definite wave-number will have an image of its own. If, then, waves of different wave-numbers are emitted from O, there will be a series of corresponding virtual images O''; and the emerging waves will not all diverge from the same point. This fact is said to be due to "chromatic aberration," because different colors have different images. (It should be noticed that the waves do not

fall normally on both surfaces; and consequently there will be spherical aberration in general, and the waves will not diverge from a *point* O'', but from a surface. See Article 317.)

It is of interest in this case to study the changes produced by the prism in a "pencil" or cone of homogeneous waves leaving O. Consider any pencil leaving O. On

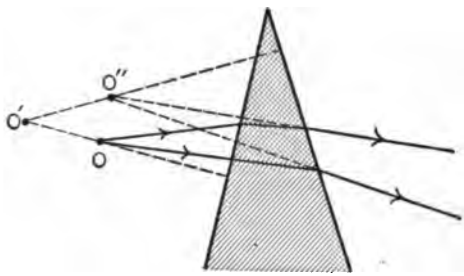


FIG. 248.

entering the prism, it will seem to come from O', and its path will be as shown. On leaving the prism, it will seem to come from O'', as shown; and it is evident that the effect of the prism is to alter the direction of the pencil. (It may be proved that, if the pencil is very small, and if it falls upon the prism so as to have minimum deviation, it will leave the prism as if it came from a *point* O''. In any other case than this the pencil will not seem to come from a point, but will be more complicated owing to spherical aberration.)

326. Plane Waves Refracted at Spherical Surfaces. The simplest case is when the surface is convex. Let S be the centre of the spherical surface; and draw from it a line $S\,C$ perpendicular to the incident plane waves. This is called the "axis." Draw an arbitrary chord for the section of the spherical surface by the paper; and let it be perpendicular to the axis, which it intersects in the point M. Then the curvature of the surface is $C = \overline{C\,M}$. When the

waves enter the surface, they cease to be plane, because in the time during which the disturbance would have gone

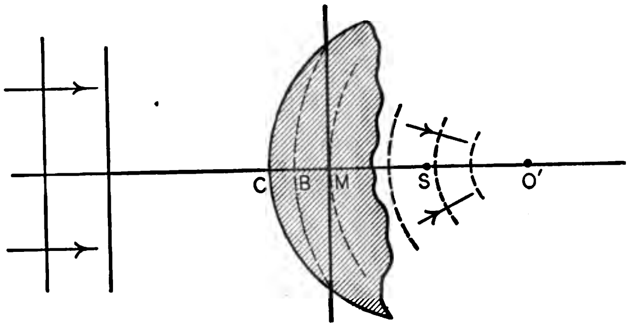


FIG. 249.

from C to M in the first medium, it goes a different distance, $\overline{CB}$, where

$$\frac{\overline{CB}}{\overline{CM}} = \frac{v_2}{v_1} = \frac{\mu_1}{\mu_2}.$$

The refracted wave-front is a sphere whose section by the paper is the circle through the point B and the ends of the fixed chord; and the waves are converging toward some point beyond the centre S. Their curvature at the surface is $\overline{BM}$. That is,

$$C_r = \overline{BM} = \overline{CM} - \overline{CB} = \overline{CM}\left(1 - \frac{\mu_1}{\mu_2}\right).$$

Hence $$C_r = C\left(1 - \frac{\mu_1}{\mu_2}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

This same formula applies equally well to plane waves incident upon a concave surface. Let the surface have a curvature $C = \overline{AM}$. When the plane waves reach it, the central disturbance along M goes to A; but at the ends of the chord of fixed length the disturbances enter the second medium a distance $\overline{MB}$, where

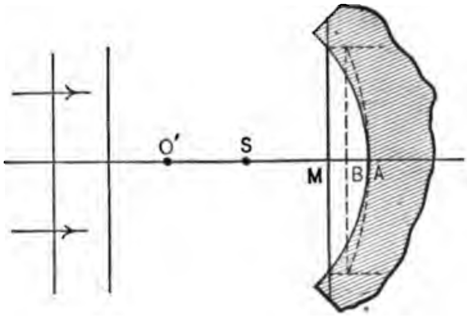


FIG. 250.

$$\frac{\overline{BM}}{\overline{AM}} = \frac{v_2}{v_1} = \frac{\mu_1}{\mu_2},$$

and the refracted wave has the curvature

$$C_r = \overline{AB} = \overline{AM} - \overline{BM} = \overline{AM}\left(1 - \frac{\mu_1}{\mu_2}\right).$$

Hence $$C_r = C\left(1 - \frac{\mu_1}{\mu_2}\right).$$

In both of these cases of refraction there is chromatic aberration if waves of different wave-numbers are incident upon the surfaces, because μ_1 and μ_2 are different for different waves, and therefore C_r varies. There will also be spherical aberration, unless only a small portion of the spherical surfaces immediately around the axis is used.

These illustrations in this article and in Article 323 are, of course, special cases of the general problem of spherical waves refracted at spherical surfaces, because a plane is but a sphere whose radius is infinite, i. e. whose curvature is zero.

327. Spherical Waves Refracted at a Spherical Surface. There are four special cases according as the waves are diverging or converging, and the surfaces concave or convex. There is one general formula which applies to all, however; and it will be deduced for the case of diverging waves refracted at a concave surface, its detailed proof for

the other cases being left as an exercise to the student, although the figures will be given.

1. *Diverging Waves Refracted at a Concave Surface.* Let S be the centre of the concave surface, and O the source of the diverging waves. Draw a line through S and O, and let it meet the surface at the point C. This line is called the axis of O. Draw an arbitrary chord perpendicular to the axis, and let it cut it in the point M. Draw a circle around the point O, so that it passes through

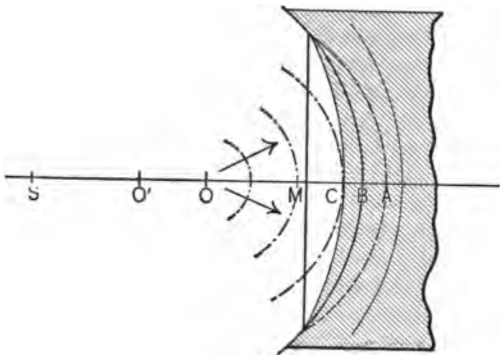


FIG. 251.

the ends of this chord; and let it cut the axis in the point A. When the disturbance from O reaches the surface at C, it advances a distance $\overline{CB}$ in the second medium in the time during which it would have advanced $\overline{CA}$ in the first medium. So the refracted wave-front passes through B and the ends of the fixed chord; and its centre is O'. Therefore, calling C, C_i, C_r the curvatures of the surface the incident wave and the refracted wave,

$$C = \overline{CM}, \quad C_i = \overline{AM}, \quad C_r = \overline{BM},$$

and
$$\frac{\overline{BC}}{\overline{AC}} = \frac{v_2}{v_1} = \frac{\mu_1}{\mu_2}.$$

Call this ratio μ_1 / μ_2, b.

By geometry
$$\overline{BM} = \overline{BC} + \overline{CM},$$
$$\overline{AM} = \overline{AC} + \overline{CM}.$$
Therefore
$$C_r = C + b\,\overline{AC},$$
$$C_i = C + \overline{AC}.$$
Hence
$$C_r = b\,C_i + C(1-b) \quad . \quad . \quad . \quad . \quad (9)$$

2. *Diverging waves refracted at a convex surface.*

$$C = \overline{CM},\quad C_i = \overline{AM},\quad C_r = \overline{BM},$$
$$\frac{\overline{BC}}{\overline{AC}} = \frac{v_2}{v_1} = \frac{\mu_1}{\mu_2} = b.$$
$$\overline{BM} = \overline{BC} + \overline{CM},$$
$$\overline{AM} = \overline{AC} + \overline{CM}.$$
Hence
$$C_r = b\,C_i + C(1-b).$$

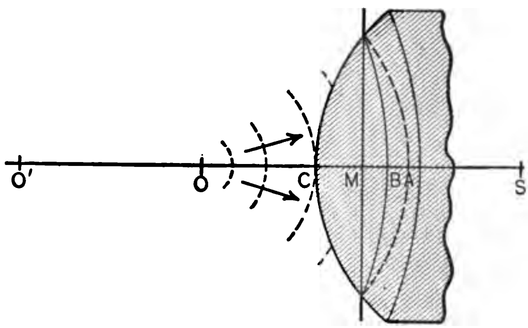


FIG. 252.

3. *Converging waves refracted at a concave surface.*

$$C = \overline{CB},\quad C_i = \overline{AB},\quad C_r = \overline{CD},$$
$$\frac{\overline{BD}}{\overline{AC}} = \frac{v_2}{v_1} = \frac{\mu_1}{\mu_2} = b.$$
$$\overline{CD} = \overline{CB} + \overline{BD},$$
$$\overline{AB} = \overline{AC} + \overline{CB}.$$
Hence
$$C_r = b\,C_i + C(1-b).$$

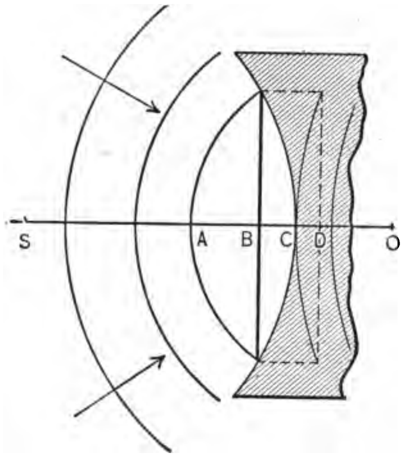


FIG. 253.

4. *Converging waves refracted at a concave surface.*

$$C = \overline{BC},\ \ C_i = \overline{AC},\ \ C_r = \overline{BD},$$

$$\frac{\overline{CD}}{\overline{AB}} = \frac{v_2}{v_1} = \frac{\mu_1}{\mu_2} = b.$$

$$\overline{BD} = \overline{BC} + \overline{CD},$$

$$\overline{AC} = \overline{BC} + \overline{AB}.$$

Hence
$$C_r = b\ C_i + C\,(1 - b).$$

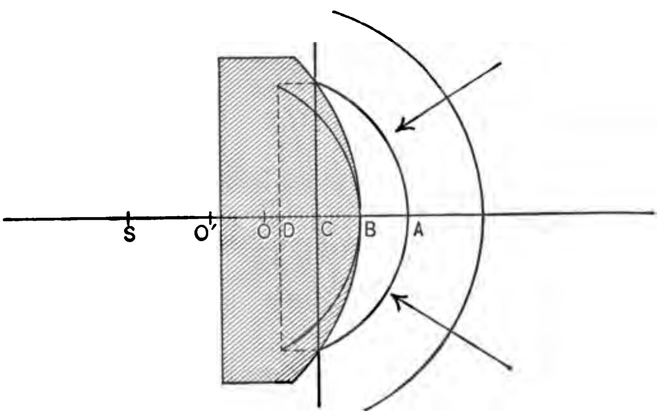


FIG. 254.

The same formula applies to all four cases, if it is remembered that the numerical values of C_i and C_r are positive if O and O' lie on the same side of the surface as S, and negative if they lie on the opposite side.

The physical description of the phenomenon is that incident spherical waves whose centre is O are changed by refraction at a spherical surface into spherical waves with their centre at O'; their curvature is changed. O' is the image of O; and it may be real or virtual.

In all these cases there will, of course, be chromatic aberration if waves of different wave-numbers are diverging from or converging to the same source; and there will be spherical aberration unless the waves are incident upon the surfaces immediately around the axis.

Lenses. A lens is a piece of transparent substance, the two principal portions of whose bounding surface are curved surfaces, usually spherical (one surface may be plane). There are many types, depending upon what kind of surfaces are combined, — concave with concave, or with convex; the concavities or convexities being in the same or opposite directions. Two cases will be treated here, — a double concave lens and a double convex one, and in both cases the lens will be considered extremely thin.

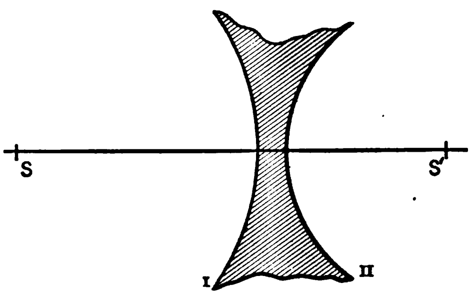


FIG. 255.

328. 1. *Double concave lens*, as shown. Let S and S' be the centres of the two spherical surfaces. The line

joining them is called the "axis" of the lens; and the point where the axis cuts the lens, if the lens is extremely thin, is called the "centre" of the lens. Consider the case of spherical waves diverging from a point O on the axis and on the same side of the lens as is S. Call the surface whose centre is S the first surface, and the other the second. Let the curvature of the first surface be C, and that of the second C'. Let the incident waves have a curvature C_i. Then, by formula (9), the curvature of the refracted waves is

$$C_r = b\,C_i + C(1 - b).$$

If the lens is very thin, the waves which enter at the first surface will reach the second before their curvature has changed; so the curvature of the waves which are incident on the second surface from within has the numerical value C_r; but its sign is negative, because a curvature which is positive for the first surface is negative for the second, since their curvatures are turned in opposite directions. Therefore the waves incident on the second surface have a curvature $-C_r$, when referred to it. In emerging from the lens into the surrounding medium the formula may be written

$$C = b'\,C_i' + C'(1 - b'),$$

in which, as just explained, $C_i' = -C_r$. Further, since $b' = \frac{\mu_2}{\mu_1}$, and $b = \frac{\mu_1}{\mu_2}$, $b' = \frac{1}{b}$. Therefore the curvature of the emerging waves is given by the formula

$$\begin{aligned} C_e &= -\frac{1}{b}C_r + C'\left(1 - \frac{1}{b}\right) \\ &= -C_i - C\left(\frac{1}{b} - 1\right) + C'\left(1 - \frac{1}{b}\right) \\ &= -C_i + (b - 1)\frac{C + C'}{b} \quad . \quad . \quad . \quad . \quad . \qquad (10) \end{aligned}$$

Now, $\dfrac{C + C'}{b}$ is an essentially positive quantity; and, if the lens has a greater refractive index than the surrounding medium, b is less than 1; and so $b - 1$ is negative. Therefore the quantity $(b-1)\dfrac{C + C'}{b}$ is an *essentially negative constant* for any definite train of waves for which μ_1 / μ_2 is constant, i. e. for a train of waves of definite wave-number. Write for it C_f. Then the above equation becomes

$$C_e + C_i = C_f \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10\,a)$$

C_e is positive if the centre of curvature of the emerging waves is on the same side of the lens as the centre of the second surface, S'. Similarly, C_i is positive if the centre of the incident waves is on the same side of the lens as S, the centre of the first surface. It must be noted that C_f is different for different values of μ_1 / μ_2, i. e. for waves of different wave-numbers; and consequently lenses have chromatic aberration.

Several special cases may be of interest. (In the figures a thick line represents the lenses, simply in order to help the clearness of the drawing.)

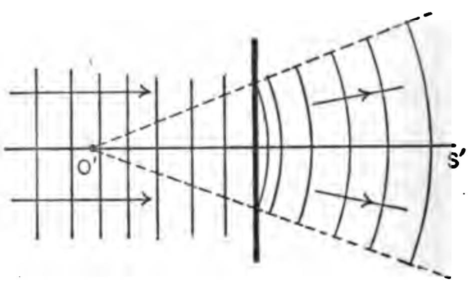


FIG. 256.

329. *a.* Incident waves are plane, and their normal is parallel to the axis of the lens; i. e. $C_i = 0 \therefore C_e = C_f$. C_f is essentially negative, and therefore the emerging waves

diverge from a point O' on the axis on the opposite side of the lens from S'. This point, O', is called the "principal focus" on that side of the lens. (Different trains of waves will have different principal foci, depending upon the value of their wave-numbers, i. e. upon μ_1 / μ_2.) The distance of this point from the lens is $1 / C_f$, and it is called the "focal distance." There is, of course, another principal focus on the opposite side of the lens at an equal distance from the lens. The effect of the lens is, then, to make plane waves diverge as if from a centre at the principal focus; and for this reason such a lens is called a "diverging lens."

A disturbance, therefore, starting from any point of a wave-front and advancing parallel to the axis will emerge from the lens as if it came from the principal focus on the opposite side.

330. *b.* Incident waves are converging on one side towards the principal focus on the opposite side; i. e. $C_i = C_f \therefore C_e = 0$. Hence the emerging waves are plane; and their wave-normal is parallel to the axis.

It follows from this that a disturbance apparently pointed toward the focus on the other side of the lens will be turned by refraction so as to emerge parallel to the axis.

This case is evidently just the reverse of case *a*, as is apparent from the diagram.

331. *c.* Corresponding to any centre of incident waves, there will, of course, be a centre of emerging waves, if the lens is bounded by surfaces of small curvatures, and if only the central portions of them are used. Therefore there will be an image of any object; and it is not difficult to construct it by graphical methods. Let F and F be the principal foci on the two sides of the lens; and let the illuminated object $O P$ be placed at O, perpendicular to the axis. P sends out spherical waves, i. e. series of disturbances in all directions. The disturbance sent out parallel to the axis will emerge as if it came from F, the

principal focus on the same side of the lens as P. The disturbance directed toward F', the other principal focus on the opposite side, emerges from the lens parallel to the axis. Therefore the image of P, i. e. the centre of all the emerging disturbances, must be the point P', the point, of intersection of the two lines as drawn. Similarly, the image of O will be at O'; and the object $O\,P$ has a virtual

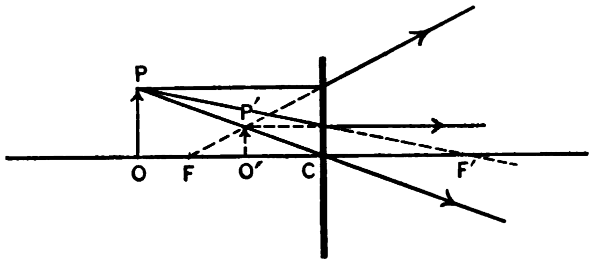


FIG. 257.

image at $O'\,P'$. The lens is supposed to be extremely thin, and a disturbance sent from P directly to the centre, C, of the lens, i. e. the point where the axis cuts it, will pass through, keeping its direction unchanged, and therefore P', the image of P, must lie on this line also.

The "linear magnification," or the ratio $\overline{O'\,P'} \,/\, \overline{O\,P}$, may be determined by means of similar triangles.

$$\frac{\overline{O'\,P'}}{\overline{O\,P}} = \frac{\overline{O'\,C}}{\overline{O\,C}} = \frac{\dfrac{1}{C_e}}{\dfrac{1}{C_i}} = \frac{C_i}{C_e}.$$

332. *d.* The general formula may be expressed in terms of distances. For, in the equation

$$C_i + C_e = C_f\,,$$

$$C_i = \frac{1}{\overline{O\,C}} \equiv \frac{1}{u}; \quad C_e = \frac{1}{\overline{O'\,C}} \equiv \frac{1}{v}; \quad C_f = \frac{1}{\overline{F\,C}} \equiv \frac{1}{f}.$$

Hence $$\frac{1}{u}+\frac{1}{v}=\frac{1}{f}; \qquad (11)$$

in which formula u is positive if the object is on the same side of surface I as its centre; v is positive if the image is on the same side of surface II as its centre; and f is an essentially negative quantity (in all ordinary lenses, such as glass surrounded by air), whose numerical value may be determined by experiment.

333. 2. *Double Convex Lens*, as shown. The centre of the first surface is now on the opposite side of the lens

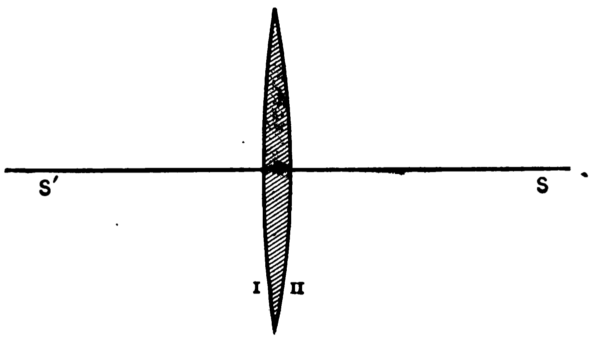


FIG. 258.

from the surface itself; and the same is true of the centre of the second surface. The formula as deduced for the double concave lens applies to this case also, because all of the equations are perfectly general, —

$$C_i + C_e = C_f;$$

where C_i is positive, if the centre of the incident waves is on the same side of the first surface as its centre; C_e is positive if the centre of the emerging waves is on the same side of the second surface as its centre; C_f is essentially negative if the lens is made of a material whose index of refraction is greater than that of the surrounding medium. In general, the centre of the incident waves will not be on

the same side of surface I as is its centre S; and it will obviously be more convenient to measure the distance of the centre of the incident waves from the surface I as *positive*, if this centre is on the *opposite* side from S, the centre of the surface I. Similarly, it is in general more convenient to measure C_e as *positive* if the centre of emerging waves is on the *opposite* side of surface II from its centre S'. Therefore, if the sign of each term in the general formula is changed, it may be written the same as before,

$$C + C_e = C_f,$$

where C_i is positive if the centre of the incident waves is on the same side of the lens as S'; C_e is positive if the centre of the emerging waves is on the same side of the lens as S; and C_f is now an essentially positive constant, if the lens has a greater refractive index than the surrounding medium, and if the waves have a definite wave-number. Its numerical value will change with the wave-number.

Some special cases may be discussed.

334. *a.* Incident waves are plane, and their wave-normal is parallel to the axis; i. e. $C_i = 0 \therefore C_e = C$. Therefore,

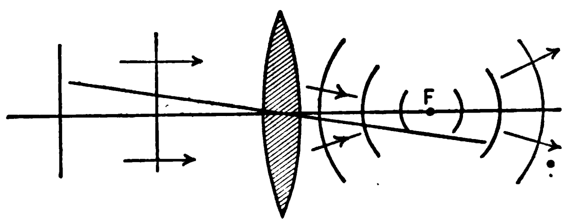


FIG. 259.

since C_f is positive, the curvature C_e is positive; and the centre of the emerging waves is at a point on the axis on the opposite side of the lens from the incident waves, at a distance from the lens $1 / C_f = f$. This point and the

corresponding one at an equal distance on the opposite side of the lens are called, as before, the principal foci; and, owing to chromatic aberration, waves of different wave-numbers will have different foci. The effect of the lens is thus seen to be to converge the waves to a focus; and so it is called a "converging lens." It follows, too, that a disturbance propagated parallel to the axis will emerge in such a direction as to pass through the principal focus.

335. *b.* Incident waves are diverging from a principal focus, i. e. $C_i = C_f \therefore C_e = o$. Hence the emerging waves are plane, with their wave-normal parallel to the axis. A disturbance, then, propagated through a principal focus and incident on the lens will emerge in a direction parallel to the axis.

This case is obviously just the reverse of case *a*, as will be seen on reference to the figure.

336. *c.* Corresponding to any centre of incident waves there is (in general) a centre of emerging waves; and

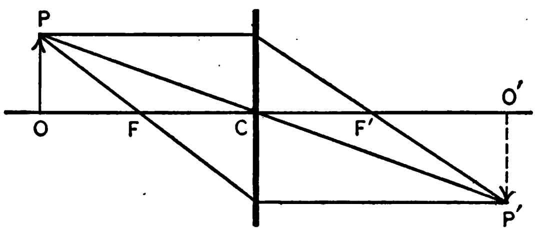


FIG. 260.

therefore there will be an image of any object. Let *F* and *F′* be the two principal foci, and let the object *O P* be placed perpendicular to the axis at a point beyond the focus *F*. The point *P* is sending out spherical waves; and the disturbance parallel to the axis will emerge from the lens in such a direction as to pass through the principal focus *F′*; the disturbance in the direction through the principal focus *F* on the same side of the lens as *P*, i. e. in the direction *PF*, will emerge from the lens par-

allel to the axis. Consequently the centre of the emerging waves is the point of intersection of these lines of disturbances, P'. The waves from P will actually converge there and then diverge again, so that P' is a real image. The lens is supposed very thin; so a disturbance from P to the centre of the lens, the point C, will pass through, keeping its direction unchanged. P' must, therefore, also lie on the line $P\,C$.

The image of O will be at O', where $O'\cdot P'$ is perpendicular to the axis; and the image of $O\,P$ is real and inverted.

The linear magnification

$$\frac{\overline{O'\,P'}}{\overline{O\,P}} = \frac{\overline{C\,O'}}{\overline{C\,O}} = \frac{C}{C_e}.$$

337. *d.* Another special case would be when the object is placed between the principal focus and the lens, the

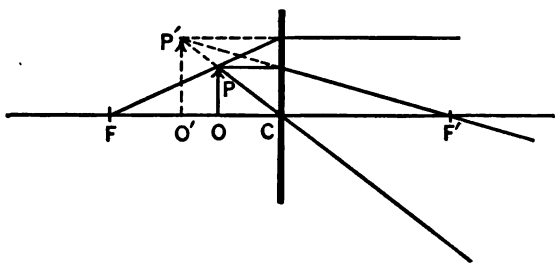


FIG. 261.

graphical solution of which is given in the figure. Spherical waves diverging from P seem to diverge from the point P' after they leave the lens. So P' is the virtual image of P. The image of the object $O\,P$ is in this position virtual and erect; it is also magnified. The linear magnification $\frac{\overline{O'\,P'}}{\overline{O\,P}} = \frac{C_i}{C_e}$ as before; and it is evident that if $C_i = C_f$, i. e. if $C_e = o$, the magnification is infinite.

338. *e.* The incident waves diverge from any point in a

plane perpendicular to the axis at the principal focus. This plane is called the "focal plane;" and, if the point O is sending out spherical waves, they will emerge from the lens as plane waves with their wave-normal parallel to the line $O\,C$, joining O to the centre of the lens. This becomes apparent if two lines of disturbances are traced, one through

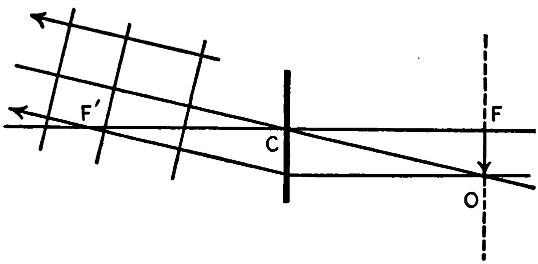


FIG. 262.

C, the other parallel to the axis and then refracted through F; for, by equal triangles, since $\overline{C\,F'}$ and $\overline{C\,F}$ are equal, the two emerging lines of disturbances are parallel. Conversely, a train of plane waves incident upon the lens will be converged to a focus in the focal plane at a point where a line through the centre of the lens, parallel to the wave-normal, meets the focal plane.

339. *f.* The general formula may also be expressed in terms of distances. For

$$C_i = \frac{1}{\overline{O\,C}} \equiv \frac{1}{u}; \quad C_e = \frac{1}{\overline{O'C}} \equiv \frac{1}{v}; \quad C = \frac{1}{\overline{C\,F}} \equiv \frac{1}{f}.$$

Hence
$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f},$$

where u is positive, if the object is on the same side of the lens as the surface which the incident waves first strike; v is positive if the image is on the same side of the lens as the surface from which the waves emerge; f is essentially positive for ordinary lenses. Stated in other words,

u is positive for diverging waves, negative for converging; v is positive for converging waves, negative for diverging.

The linear magnification obviously equals v / u.

Combinations of Lenses. Various combinations may be arranged; but only a few are of fundamental importance.

340. 1. *Microscopes.* A microscope is an instrument designed to magnify the apparent size of objects. The simplest form is the ordinary "magnifier," which consists of a single converging lens (or converging system), which is so placed that the object comes just inside the focus. Then, as explained in Article 337, a magnified image of the object will be formed, which may be seen by the eye of an observer.

A compound microscope, as commonly used by microscopists, consists of two converging lenses L_1 and L_2, which

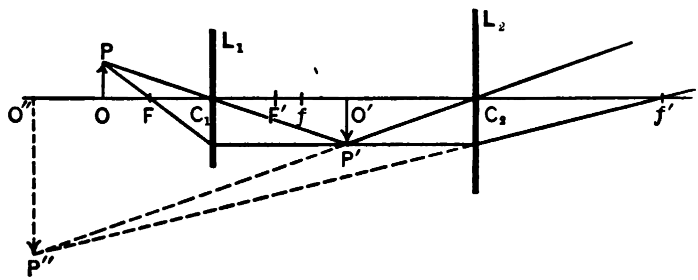


FIG. 268.

can be so adjusted that, when the object to be examined is placed outside the focus of L_1, its image will fall just inside the focus of L_2. Consequently, there will be a real image formed by L_1; and this will be magnified by L_2 into a virtual image. The first lens is called the "objective;" the second, the "eye-piece." In the figure $O\,P$ is the object; $O'\,P'$ is the real image formed by L_1; $O''\,P''$ is the virtual magnified image formed by L_2.

The linear magnification is $\dfrac{\overline{O''\,P''}}{\overline{O\,P}}$. But

$$\frac{\overline{O'' P''}}{\overline{O' P'}} = \frac{\overline{O'' C_2}}{\overline{O' C_2}}; \quad \text{and} \quad \frac{\overline{O' P'}}{\overline{O P}} = \frac{\overline{O' C_1}}{\overline{O C_1}}; \quad \text{and}$$

therefore
$$\frac{\overline{O'' P''}}{\overline{O P}} = \frac{\overline{O'' C_2} \times \overline{O' C_1}}{\overline{O' C_2} \times \overline{O C_1}}.$$

In actual practice $\overline{O C_1}$, the distance from the object to the objective, is made as small as possible, consistent with its being beyond the focus; and $\overline{O'' C_2}$ has a limiting value depending upon the distance at which the eye of an observer can see the image clearest. In most cases this distance is about fourteen inches.

The instrument must be so designed also that the cone of waves produced by any point of the object is slightly

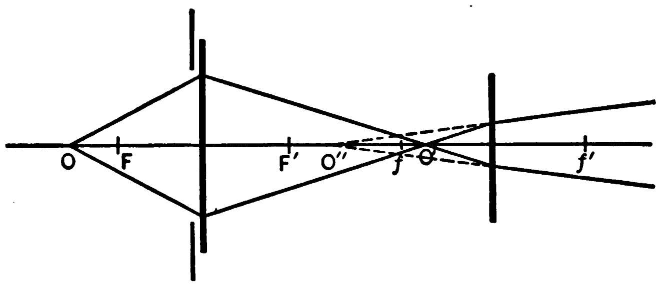


FIG. 264.

larger than the aperture of the pupil of the eye, when the waves finally emerge from the last lens. This is illustrated in the figure. The size of the objective L_1 is always limited by some diaphragm or stop; and so the cone of waves which start from O will have the path as shown; the final emerging cone of waves seeming to come from O''. In practice, diaphragms are introduced in between the two lenses so as to cut off all the portions of the lenses except near the axis.

It should be noted that this process of magnification cannot continue indefinitely; for it may be proved that if two points are as close together as $\lambda / 2$, half the wave-

length of the light used, no magnification can "resolve" them, that is, show them as two distinct points.

341. 2. *Telescopes.* A telescope is an instrument designed to render distant objects more clearly visible to the eye. There are a great many types of telescopes, some depending upon the use of concave mirrors, others upon lenses. The former are called "reflectors;" the latter, "refractors." Refractors are of two types, drawings for both of which are given. The first consists of two converging lenses L_1 and L_2 so placed that their two foci, F' and f, almost coincide.

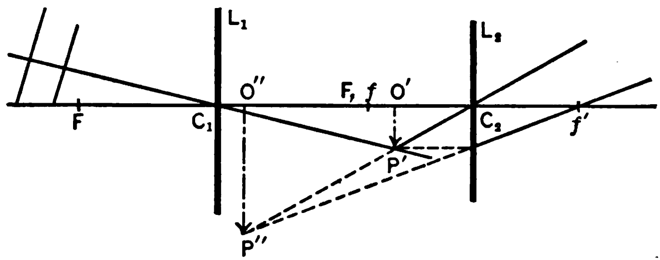


FIG. 265.

The waves coming from a distant object and falling upon L_1 are brought to a focus at $O'\,P'$ just beyond the focus F'. For the waves converging to P' are coming from a point P in the line $P'\,C_1$ so far away that the waves which it emits are nearly plane when they reach the lens; and, similarly, O' is the focus of nearly plane waves emitted by a source O, very far away on the line $O'\,C_1$. This real image lies just inside the focus f of the lens L_2; and so a magnified virtual image will be formed. The power of the telescope must be expressed as the ratio of the apparent size of this virtual image $O''\,P''$ to the apparent size of the original object if viewed with the naked eye; and these sizes may best be compared by means of the angles which are subtended at the eye in the two cases. The angle subtended by $O''\,P''$ is $O''\,C_2\,P''$; that subtended by the distant object is $O'\,C_1\,P'$. Therefore, the power of the instrument is

$\frac{O'' C_2 P''}{O' C_1 P'} = \frac{\overline{C_1 O'}}{\overline{C_2 O'}}$ if the angles are small. But $\overline{C_1 O'}$ is the focal length of L_1; and $\overline{C_2 O'}$ is the focal length of L_2 if F' and f coincide and the object is very far away. If these two focal lengths are F_1 and F_2, the power of the telescope is F_1/F_2. The first lens is sometimes called the "object-glass;" and the second, the "eye-piece." The focal length of the object-glass must be large so as to have great power; and the lens itself must also be as large as the nature of the material will permit, in order to receive as much light as possible from the distant object.

342. The second type of refracting telescope consists of two lenses L_1 and L_2; the one nearer the distant object,

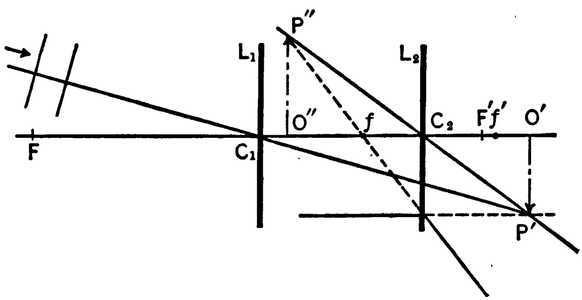


FIG. 266.

L_1, is converging; the second, L_2, is diverging. They are so placed that the two foci F' and f' nearly coincide, as shown. The converging lens would form an image of the distant object at $O' P'$; but, owing to the presence of the diverging lens, the waves converging toward P' are so changed as to become diverging from a virtual image P''. Similarly, the waves converging toward O' will be diverged by L_2 until they seem to come from O''. So that a virtual, erect image $O'' P''$ is formed of the distant object. This is the construction and principle of the ordinary "opera-glass."

343. 3. The human eye is, in one respect, an optical instrument, because it consists of a combination of lenses which focus upon the retina light-waves coming from illuminated objects. The eye has but little spherical aberration, owing to its peculiar shape and to the action of the iris, which takes the place of a diaphragm; but it does have considerable chromatic aberration. Most eyes have power of "accommodation," that is, of altering their focal length at will so as to perceive objects at different distances away. There are, however, several possible optical defects in eyes, which may arise from various causes: —

a. The waves may be brought to a focus in front of the retina instead of on it. Such eyes are called "near-sighted," and may be helped by the use of diverging lenses.

b. The waves may be focused back of the retina. Such eyes are called "far-sighted," and may be helped by the use of converging lenses.

c. The focus may be different for different sections of the eye. For instance, if the dial of a clock is looked at, an eye may see the figures 2 and 8 clearly, but may not see the 5 and 11 sharply. Such eyes are called "astigmatic," and may be helped by the use of cylindrical lenses.

d. In normal eyes, the images formed by the two eyes of the same object fall upon two points in the retinas which are said to "correspond;" and only one visual impression is made. But it may happen, owing to muscular troubles, that the images formed by the two eyes do not correspond; and so two visual impressions are seen of the same object. Such eyes may be helped by the use of prisms.

CHAPTER IV

DISPERSION — SPECTRA

344. Dispersion. It has been shown in the last chapter that, when a train of plane waves passes through a prism, the direction of the wave-normal is changed; that is, there is "deviation." This deviation was shown to depend upon several conditions, — the wave-number of the incident waves, the angle of incidence, the material of the prism, and the angle of the prism. It was further proved that waves of greater wave-number (i. e. of less wave-length) are (in general) deviated more than those of a smaller wave-number. If, then, two trains of waves of different wave-number fall upon the same prism at the same angle of incidence, they will be deviated differently; and the difference in the angles of deviation, that is, the angle between the normals of the two emerging waves, is called their "dispersion." As might be expected, the dispersion of the same two trains of waves is different for different angles of incidence and for different prisms; and, for the same angle of incidence and the same prism, the dispersion is different for different trains of waves, even if they have the same difference in their wave-numbers. Since, when two trains of plane-waves in the same direction but of different wave-numbers are refracted through a prism, they emerge in different directions, it follows that, if these same two trains of waves are made to fall upon the prism at the same angles as those at which they emerged, they will be refracted out so as to emerge with their wave-normals in the same direction.

345. Pure Spectrum. If waves coming from a source which appears white to our eyes are made to pass through a prism and then fall on a screen, it is observed that the illumination on the screen is not white, but colored. If the source of light is a point, there will be, as explained in Article 325, a series of virtual images corresponding to each train of waves of a definite wave-number, which is emitted by the source; and each one of these images will be the centre of a train of diverging waves, which will illuminate the screen. There will, therefore, be an overlap-

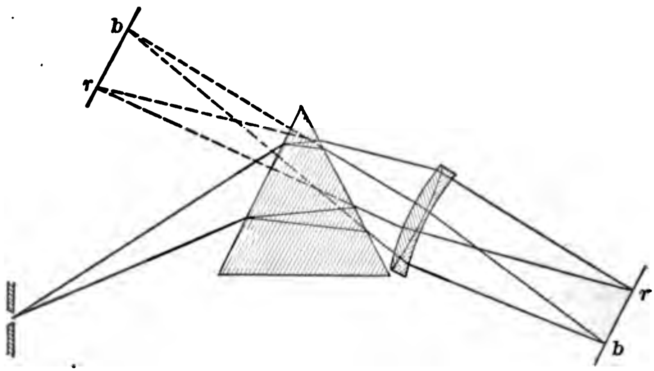


FIG. 267.

ping of the effects due to the separate waves. This may be prevented by interposing between the virtual images and the screen a converging lens, which will focus on the screen the waves diverging from the images. There will be, under these conditions, a separation of the waves, so that, corresponding to each virtual image of the source, there is a definite point of illumination on the screen. All the different trains of waves emitted by the source will thus be separated and spread out, as it were, on the screen, each illuminated point receiving only waves of one wave-number. In practice, the source of light is not a point, but is a narrow slit placed parallel to the edge of the prism, and in front of the flame or light; and the images

on the screen are, then, not points but narrow lines parallel to the slit. When the waves from any source are thus resolved into their components by a slit, a prism and a lens, the resulting single waves are said to form the "spectrum" of the source. If the slit is extremely narrow, its images on the screen will be narrow; and there will not be much overlapping of the separate images due to the separate trains of waves. Such a spectrum is called "pure" to distinguish it from one where there is overlapping, produced either by width of slit or absence of lens. The purest spectrum is obtained when the incident waves fall upon the prism as nearly as possible at the angle which corresponds to minimum deviation; for, under these conditions, the virtual images of the slit are the sharpest possible. (See the end of Art. 325.)

If the source is one of ordinary white light, such as a white-hot solid, and if the prism is one of ordinary glass, its spectrum, as shown on the screen, will be a series of waves which produce in our eyes the sensations blue, green, yellow, and red, as well as all intermediate shades; and so "white light" may be regarded as a mixture of all those waves which produce these individual sensations. Those waves which produce the sensation blue will, as noted several times, be deviated more than those which produce the sensations green and red, if the prism is any ordinary one; and the arrangement of the colors in the spectrum will be in the order of the wave-numbers of the corresponding waves. In the spectrum of the white-hot solid there are, of course, other waves than those which appeal to our sense of vision; some are longer than the "visible waves," and may be detected by their heating effect, — they are called the "ultra-red waves;" others are shorter than the visible waves, and may be detected by their action on certain chemicals such as a photographic plate, — they are called the "ultra-violet waves."

346. Anomalous Dispersion. There are certain prisms

which do not disperse the waves in the order of their wave-numbers; and such dispersion is called "anomalous." If waves from a white-hot solid pass through such a prism, there is always a gap in the spectrum, perhaps several, showing that certain waves have not come through but have been "absorbed" by the prism; and the waves on each side of a gap are displaced, so that those waves which are longer than the ones absorbed are deviated more than they ordinarily would be, and those which are shorter are deviated less. As seen on a screen, or as viewed by the naked eye, the colors on each side of the absorption band are shifted towards and over each other. There is a most intimate connection between this phenomenon and certain others to be described later. (See Art. 355.)

An illustration of anomalous dispersion is afforded by a prism of alcoholic solution of fuchsine (aniline-red). The greenish-blue is absorbed; and the order of the colors, according to deviation, is violet, blue, red, yellow.

347. Achromatism. As has been repeatedly mentioned, a prism or a lens has chromatic aberration; that is, the image of any source of light as formed by it depends upon the wave-number of the waves; and, if a source of white light is used, there will be a series of colored images. This is, of course, most inconvenient for many purposes; for instance in the method for securing a pure spectrum, an ordinary lens would not focus all the waves at the same distance from the lens. It is possible, however, to make combinations of lenses which will largely obviate this difficulty.

As stated in Article 344, it may be proved by experiment that the dispersion of any two trains of waves depends upon the prism, its material and angle, and upon the angle of incidence. Therefore it may be possible to make two prisms of different material and angle, such that they will produce the same dispersion of these two trains of plane waves; although, to produce this, the angle of inci-

dence upon the two prisms must, in general, be different, and the angles of emergence will be different. If these two trains of plane waves are incident together upon one prism, they will emerge dispersed through a certain angle; and, if they now fall upon the other prism which is so placed as to refract them together, they will emerge from it with their wave-normals parallel, as shown; for one prism simply neutralizes the dispersion of the other. The

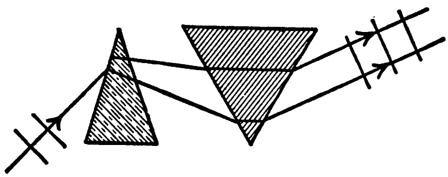


FIG. 268.

direction of the wave-normal of the emerging waves is not, in general, the same as that of the incident waves; and, therefore, there is deviation but no dispersion. By a suitable choice of prisms, it is possible, of course, to prevent the dispersion of two trains of waves of any wave-numbers; but, if in the incident waves there are some of other wave-numbers, all these others will be dispersed slightly, although not so much as ordinarily with a single prism. Thus, if waves from a "white-hot" body are falling on the pair of prisms, two trains of waves will emerge in the same direction; and the others, in slightly different directions.

In a perfectly similar manner it is possible to combine two lenses of different material and curvatures, one converging, the other diverging, so that the combination brings to the same focus any two trains of waves. This is possible because, as shown for the two prisms, there is, in general, deviation produced by the combination; and, therefore, so far as the ordinary laws of prisms and lenses are concerned, the combination of the two produce the same

effect as one if suitably chosen. Consequently, plane waves will be brought to a definite focus, etc. Such a combination is called an "achromatic lens," and is said to be "corrected for" those two definite wave-numbers. The other waves will not be brought to the same focus, but will not deviate far from it. If a lens is to be used for visual work, it is generally corrected for the two strongest colors in the visible spectrum which are some distance apart, perhaps the blue and yellow. Then these corresponding waves will be brought to the same focus, and the other waves will come nearly to the same point; so that, if waves from a white-hot body pass through such a lens, the light at the focus will be nearly white, although it still is slightly colored. These remaining colors are sometimes called "secondary colors." A lens which is to be used for photographic work is, naturally, corrected for two waves which produce great photographic action.

348. Direct-Vision Spectroscope. Two prisms of different material and angle may, of course, produce the same deviation for any one train of waves; but, since their dispersion will be different, the deviation of any other train of waves will not be the same in both prisms. Therefore, if these two prisms are combined so that one tends to neutralize the deviation produced by the other, that is, if they are placed each with its edge near the base of the other, they will form a system such that, if the waves from a white-hot body pass through them, one particular train of waves may not be deviated at all, but will emerge parallel to its original direction, while the others will be slightly deviated from this direction. Consequently there will be dispersion. If several prisms are used, considerable dispersion may be secured; and by suitable arrangements of a slit and lens a pure spectrum of the source will be obtained. Such an instrument is called a "direct-vision spectroscope," because the entire apparatus for viewing the spectrum is all arranged in one straight line.

349. Ordinary Spectroscope. The ordinary spectroscope, or instrument for the study of the spectra emitted by different sources, is a simple modification of the apparatus described before, in Article 345, for producing a pure spectrum. It consists of three separate parts: the collimator, the prism (or train of prisms), and the telescope. The collimator is a metal tube at one end of which is placed the slit, and at the other an achromatic converging lens, the tube being of such a length that the slit is at the principal focus of the lens. It is carried by a rigid arm which can turn around an axis perpendicular to the

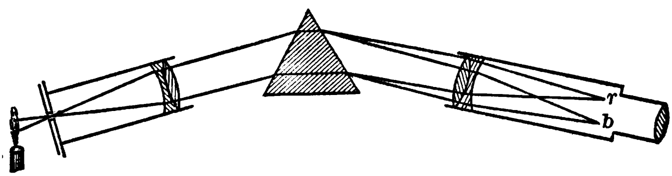


FIG. 269.

platform which carries the prisms. This axis is called the "axis of the instrument;" and the collimator is adjusted perpendicular to it. Several prisms are generally used, so as to produce greater dispersion; and the different prisms are all placed on a platform with their edges parallel to the axis of the instrument, and on the same level as the collimator. They are, further, all joined by such a link-work that, however the collimator is turned around the axis of the instrument, the prisms are so moved that waves from the collimator will fall upon the first prism and upon the others in turn, at the angle which corresponds to minimum deviation. The telescope is an ordinary instrument with two converging lenses, as explained in Article 341. It is carried by a rigid arm which is movable about the axis of the instrument; and it is adjusted perpendicular to this axis, and at the same level as the collimator.

If, now, the slit is made very narrow, and is illuminated by any source, the corresponding waves will leave the lens of the collimator with plane wave-fronts, will fall upon the prisms at the angle which corresponds to minimum deviation, will leave the prisms as if coming from a series of virtual images at an infinite distance, and will be focused by the object-glass of the telescope into a series of lines in the focal plane, each line corresponding to a train of waves of a definite wave-number. This series of images is then magnified by the eye-piece of the telescope. In this manner the nature of any source of light may be studied.

It is convenient, of course, to have some means of locating or distinguishing the different trains of waves and their images; and one of two methods is commonly adopted: one is to throw into the focal plane of the telescope a fixed scale divided into convenient lengths; the other is to illuminate part of the slit with some standard source, and to compare with its spectrum the waves coming through the other part of the slit, due to the source to be investigated.

350. Continuous and Discontinuous Spectra. When the waves coming from various sources are thus analyzed by means of a spectroscope, it is observed that there are two classes of spectra, — one where all possible waves seem to be present, the other where there are certain gaps due to the absence of particular waves. The former is called a "continuous" spectrum; and it is proved by experiment that a solid body emits such a one. That is, if a solid is white-hot, it emits all the waves which lie in the visible spectrum, as well as others, both longer and shorter, there being no waves absent between the limiting wave-lengths. The other kind of a spectrum is called "discontinuous;" and it is emitted by vapors and gases, when they are excited to luminescence in any way, e. g. by passing an electric spark through them.

This difference between solids and gases is what might

be expected on the kinetic theory of matter. In gases the molecules have certain "free paths;" and while the molecules are moving along them, their parts may vibrate freely in definite periods, thus producing trains of waves of definite wave-numbers. In a solid, however, the molecules are so close together, and so connected, that any definite periodic vibration is impossible, and all possible periods and corresponding waves occur.

The spectrum of any gas or vapor is perfectly characteristic of itself; it consists of many isolated waves which are peculiar to itself. So, by examining the spectrum of any gaseous source, it is often possible to detect what gases or vapors are present there.

351. Emission and Absorption Spectra. The waves which are emitted by any source, such as an incandescent gas, are said to give an "emission" spectrum when analyzed by a spectroscope.

If waves from a white-hot solid are made to enter any substance, the emerging waves may be different from those entering; some waves may have been "absorbed." This may be tested by analyzing the emerging waves by means of a spectroscope. The white-hot solid emits a continuous spectrum; and, consequently, any gaps in the spectrum must be due to the absorption of the corresponding waves by the substance interposed between the slit of the spectroscope and the white-hot source. These waves which are absorbed by any substance are called the "absorption spectrum" of that substance. In general, the absorption spectrum of a liquid consists of whole groups of waves; and so there are wide gaps in the spectrum as seen in the spectroscope. But the absorption spectrum of a vapor consists of separate isolated waves.

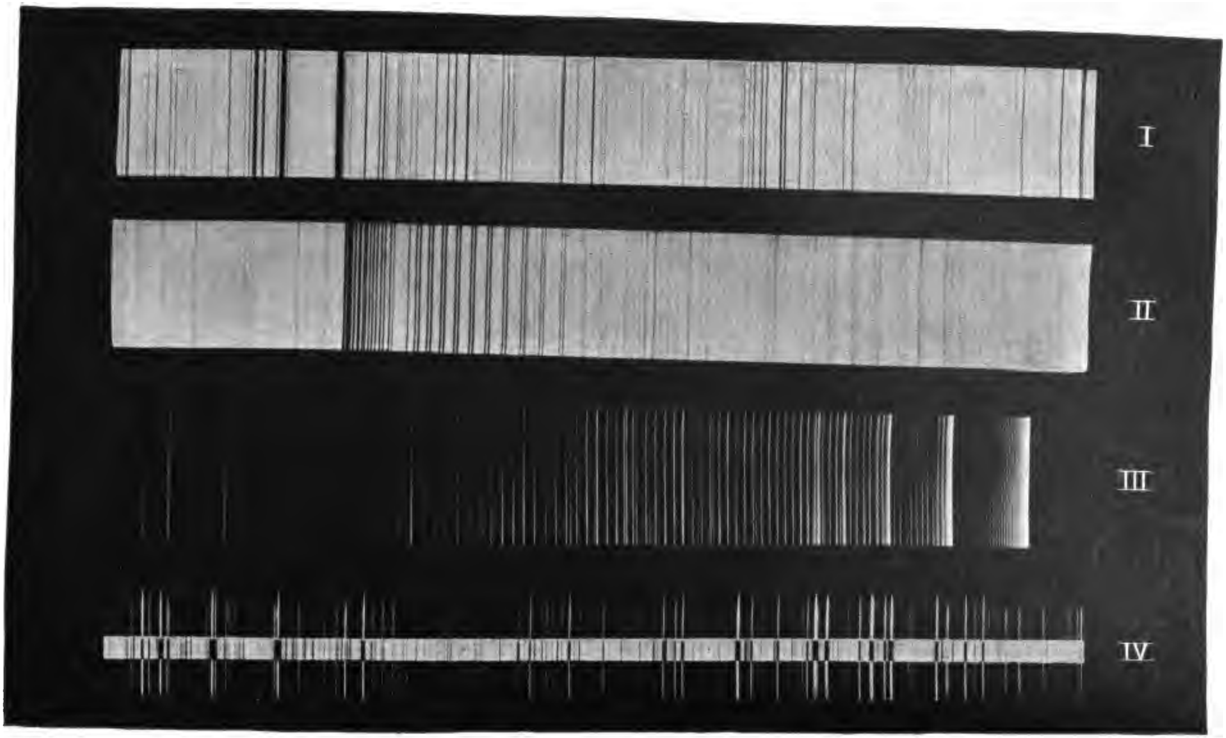
There is a most interesting connection between the emission spectrum and absorption spectrum of the same vapor at the same temperature; in fact, they are identical. That is, those waves which a vapor emits are exactly those

which it will absorb out of all the waves of a continuous spectrum. This is but a special application of the general law of radiation and absorption, which was explained in HEAT (Art. 206).

The intensity of the emission or the absorption depends, of course, upon the temperature of the vapor; the higher the temperature, so much the more intense is the emission.

352. Fraunhofer's Lines. The spectrum of the waves coming from the sun is found on examination to be an absorption one; and the explanation is obvious: the sun consists of a white-hot central mass which emits a continuous spectrum, and which is surrounded by layers of absorbing vapors. These vapors, of course, emit definite waves; but, as their temperature is so much lower than that of the central mass, those waves which come through from the inside, that is, those which are unabsorbed, are so intense, compared with those emitted by the outer layers, that the latter produce no effect comparable with the former. When sunlight is analyzed, then, by means of a spectroscope, there will be in the spectrum dark lines, which are parallel to the slit, and which mark the absorbed waves. These lines are called "Fraunhofer's Lines," from the name of the German astronomer who first carefully studied them.

Since the waves absorbed by any vapor are exactly those which it emits, it is seen that, if it can be proved that certain vapors emit waves whose absorption lines are in the solar spectrum, that fact demonstrates the existence of those vapors in the layers around the central mass of the sun. The emission spectra of a great many vapors have been carefully examined, and compared with the solar spectrum; and almost all the Fraunhofer lines have been accounted for. The following substances undoubtedly exist on the sun: calcium, iron, hydrogen, sodium, nickel, magnesium, cobalt, silicon, aluminium, titanium, chromium, manganese, carbon, and perhaps thirty others. Antimony,



DESCRIPTION OF PLATE

I. Portion of Solar Spectrum, showing absorption lines due to metallic vapors on the Sun.

II. Portion of Solar Spectrum, showing a series of absorption lines due to the presence of oxygen in the earth's atmosphere.

III. Portion of the Emission Spectrum of Carbon.

IV. Portions of the Solar Spectrum and the Emission Spectrum of Iron-vapor, showing by the coincidences the existence of iron-vapor in the atmosphere of the Sun.

arsenic, bismuth, gold, boron, mercury, and some others do not occur so far as is known.

In a similar manner the nature of the substances on many of the stars has been studied; for in some cases their spectra are emission, in others absorption.

353. Absorption. When waves from a white-hot solid fall upon any substance, it may happen that some waves of particular wave-numbers do not enter the surface; others that do enter may not pass through to the further side; and, consequently, of those waves which are thus absent from the emerging trains, some are "absorbed" at the surface, others in the interior of the substance. (This surface-effect has no connection with ordinary reflection, for in that all waves are reflected alike.)

Absorption implies the removal from the entering waves of some particular trains; and, as waves carry energy, an explanation must be given of the way in which this absorbed energy becomes manifest.

354. *a. Surface-absorption.* An illustration of this is given by any metallic surface, or by a surface of almost any of the aniline dyes. If waves from a source of white light fall upon such a surface, certain definite waves are reflected at once, apart from a certain percentage of all the waves which is reflected owing to ordinary reflection. Thus a surface of gold reflects those waves which produce in our eyes the sensation yellow. The other waves enter the gold and are soon absorbed there, unless the gold is in the form of a film.

If a sufficiently thin film of a metal is made, it is transparent to certain waves; and it is possible also to make prisms of most metals. It is observed that such prisms have "anomalous dispersion" (see Art. 346). A prism made of a solution of an aniline dye also has anomalous dispersion.

355. *b. Ordinary absorption.* As trains of waves pass through any medium, certain waves are as a rule absorbed;

and their energy may be spent in several ways. In general, it is used in producing "heat-effects" in the absorbing medium. Other waves are not absorbed, and can pass through the medium freely.

Thus, if waves from a source of white light enter a leaf of a tree, certain waves are absorbed, and others pass through, or are reflected out from the interior if they strike any reflecting bodies inside. The waves which emerge produce color sensations in the eye of an observer; green, in the case of a leaf. This is the explanation of the colors of most natural objects.

The amount of the absorption of any particular train of waves depends upon the distance through which the waves pass; and so it often happens that thin and thick layers of the same substance have different colors, if it absorbs the two corresponding trains of waves differently.

356. *c. Fluorescence.* In certain substances (perhaps all), some of the energy received by the absorption of the waves is not spent in heat-effects, but in emitting new waves. Such substances are called "fluorescent." It is a general law that the wave-length of the waves absorbed is always shorter than that of the waves emitted by means of the corresponding energy. Thus "canary glass," a glass containing uranium oxide, absorbs the blue waves and emits greenish-yellow.

357. *d. Phosphorescence.* Most fluorescent bodies emit the fluorescent waves only while they are absorbing other waves; but certain of them continue to emit waves for some time after absorption has ceased. Such bodies are called "phosphorescent." They store up the absorbed energy, and emit it slowly. Many sulphides are phosphorescent; and, if they are exposed to sunlight for a few minutes, and are then carried into a dark room, they will continue to shine for many hours.

358. *e. Reflection by Fine Particles.* If a cloud of fine particles is interposed in a beam of white light, it may be

observed that, when looked at from the side, the cloud appears colored blue. This is due to the reflection and scattering by the fine particles of the shortest waves, those which produce in the eye the sensation blue. The other waves pass by or around the particles; for a body can reflect only those waves which are considerably smaller than its own dimensions. This is the explanation of the color of the blue sky.

CHAPTER V

COLORS — COLOR-SENSATION

359. Complementary Colors. It has been shown that, when the waves from a white-hot solid are analyzed by a spectroscope, there is a continuous spectrum, containing waves of all possible wave-numbers between certain limits. Corresponding to any one of these wave-numbers which is in the visible spectrum, there is a definite sensation of color. So white light is commonly said to be a mixture of a great many colors. If one of these trains of waves is absent from the white light, the other waves will combine to produce in the eye of an observer a certain color-sensation, which is called the "complementary" color to that which the absent train would produce. Thus, red and greenish-blue are complementary colors; so are yellow and blue.

Conversely, if these two sensations are produced simultaneously in the eye, it will receive the sensation of white light. (This has nothing at all to do with the color-effect of mixing paints or colors; for in that there is simply double absorption of the waves, and the color-sensation produced corresponds to those waves which are not absorbed by either paint.)

In a similar manner, if the waves which remain after one train is removed are separated into two sections, each section will produce a definite color-sensation; and white light may be said to be resolved into three colors. It may be proved by experiment that, using different intensities

of any three colors which together compose white light, it is possible to produce any color-sensation desired.

The simplest method of combining color-sensations is to paint sectors of a circular disc with the different colors, and then rapidly to revolve the disc in its own plane. Under these conditions the eye receives practically simultaneous sensations from all the colors.

360. Colors of Objects. The colors of objects depend upon three things, — the nature of the light in which they are viewed, the absorption of the object, the peculiarities of the eye of the observer.

An object is ordinarily called white if it reflects all the visible waves; and so, if such an object, like a piece of paper, receives only waves which produce the sensation yellow, the object will appear yellow. The names of the colors associated with all natural objects depend upon their illumination by white light. A stick painted red will appear black if placed in any colored light except red, unless in producing this red-colored paint some other color has been mixed in; in which case, if placed in light of that color, the stick will appear to be of the same color.

There are as many kinds of absorption-colors as there are modes of absorption. A film of gold-leaf appears yellow if white light is reflected from it, owing to surface-absorption; and the complementary waves will pass through so that the transmitted light has the color blue. As previously explained, the colors of almost all natural objects depend simply upon the withdrawal of one or more colors by absorption in the interior of a body. In certain cases, moreover, the effect is complicated by the presence of fluorescent colors.

The normal eye is able to distinguish a continuous gradation of colors from the darkest violet to the deepest red, and to distinguish some colors which do not appear in the spectra of bodies, such as purple and brown. But there are people so unfortunate as not to be able to

distinguish all these colors; and such are called "color-blind." It happens, not unfrequently, that people cannot distinguish red from green; and therefore, in the case of all colors, such people would have peculiar and individual sensations.

361. Perception of Color. In a normal eye it is possible to distinguish two entirely different sensations, — one of gray, and one of color. Any train of waves which, if intense, would produce a color-sensation, will, if very weak, produce the gray sensation; and again there are portions of the retina of the eye where only gray can be perceived.

It is now a recognized fact that the sensation of color is a differentiation of the gray sensation; and that, in all probability, different structures of the retina have particular functions to perform in these perceptions. It seems probable that the so-called "cones" of the retina are the instruments of color-sensation; and that the "rods" are for the gray sensation.

CHAPTER VI

INTERFERENCE — DIFFRACTION

362. Interference Due to Two Sources. In Young's interference experiment, as described in Chapter I., Article 300, two slits in a screen were illuminated by light from a parallel slit in another screen, and the slits were so arranged that the two slits received identically the same illumination. Under these conditions there were dark and bright bands on the screen, if the light was homogeneous, that is, if the train of waves had a definite wave-number. It was absolutely essential that the two slits should have the same illumination in every way; that is, if there is any change in one, there is the same change in the other, etc. For, if two different sources of light are used, e. g. if two candle flames are placed behind the two slits, there will no longer be interference. Any source of light emits the waves irregularly; and, if there is to be complete interference between waves coming from two sources, it is obvious that the two sources must undergo identically the same changes.

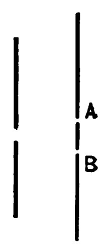


FIG. 270.

To secure interference, then, between waves from two sources, these two must be identical. Various methods have been devised to satisfy this essential condition.

a. The simplest apparatus is, as already described, a slit on one screen, and two parallel slits in a second screen parallel to the first; the second screen being so placed that the opaque portion between the two slits is exactly opposite the first slit. If a

source of homogeneous waves is placed in front of the first slit, the two slits will receive identical illumination, and so be identical sources.

b. A slit is placed parallel to the edge of a "biprism." A biprism, of which a section is shown, consists, as it were, of two very thin prisms, exactly alike, joined at their bases. The slit is placed opposite the thickest part of the biprism; and waves from any source in front the slit will fall upon the prism, and will emerge as if coming from two virtual slits at *A* and *B*, as shown. These two virtual images are identical; and so are the trains of waves which apparently come from them.

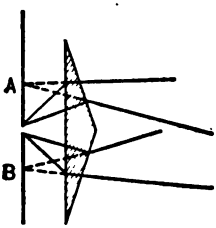


FIG. 271.

c. Two plane mirrors are inclined to each other at a most obtuse angle, and a slit is placed so as to illuminate both. There will be an image of the slit in each mirror; and the two images, *A* and *B*, will be identical. The adjustments for this experiment are quite difficult.

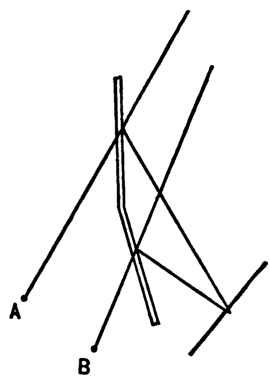


FIG. 272.

d. A slit is adjusted parallel to a plane mirror, but slightly above it. Any screen placed suitably will receive now two trains of waves; one from the slit directly, the other from the virtual image of the slit in the mirror.

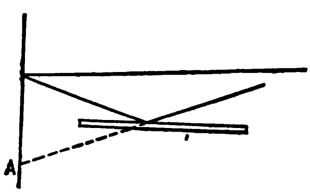


FIG. 273.

In all these four methods care must be taken to have the two sources close together and accurately parallel; under these conditions a source of

homogeneous waves in front of the first slit will produce light and dark interference bands on a screen which is parallel to the plane of the two sources.

As stated in the first chapter, there is a connection between the distance apart of the bands on the screen and the wave-length of the waves. This may be best shown by deducing the exact mathematical connection.

Let A and B be the two identical sources of light, which are supposed to be very close together. Let the screen be parallel to the two slits A and B, and placed at a considerable distance away. Draw a perpendicular line from the opaque portion between the slits to the screen, and let it meet it in the point C. A and B are so close together, and C is so far away, that the lines $A\,C$ and $B\,C$ may be regarded as having the same length as the perpendicular line, and as being parallel. Call the distances $A\,B$, c, and $A\,C$ (or $B\,C$), a. The disturbance at any point, P, on the screen is the resultant of two disturbances which come from the two slits A and B. Their difference in path may be found by dropping a perpendicular, $A\,D$, from A upon the line joining B and P. Call the distance from C to P, $C\,P$, $= x$; then by similar triangles

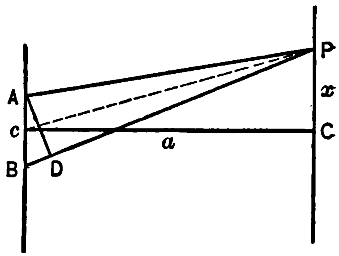


FIG. 274.

$$A\,D : A\,B = P\,C : A\,C,$$

or the difference in path, $A\,D = \dfrac{c\,x}{a}$.

If this difference in path is an odd number of half wave-length, there will be no resulting disturbance at P; whereas, if it equals a whole number of wave-lengths, the effects will strengthen each other, and there will be a bright band.

1. *Dark bands.*

The condition is that $\frac{cx}{a} = (2n + 1)\frac{\lambda}{2}$; where n equals any whole number 0, 1, 2, 3, etc. Corresponding to any value of n there will be a definite value of x; and therefore there will be a succession of dark bands across the screen, each band being parallel to the slits.

The distance apart of two bands is found by giving n two consecutive values, b and $b + 1$. Thus the distance equals $\frac{a\lambda}{c2}\{2(b+1) + 1 - (2b + 1)\}$. Calling it D,

$$D = \frac{a\lambda}{c} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

2. *Bright bands.*

The condition is that $\frac{cx}{a} = n\lambda$,

where n equals in turn 0, 1, 2, 3, etc. Where $n = 0$, i. e. where $x = 0$, there is a bright band for all values of λ. All waves, therefore, no matter what their wave-numbers, give a bright band at the centre of the screen, directly opposite the opaque portion between the two slits.

The distance apart of two consecutive bright bands is, as before for dark bands,

$$D = \frac{a}{c}\lambda\{(b+1) - b\} = \frac{a\lambda}{c}.$$

Consequently the bright and dark bands are equally spaced so long as the wave-length is constant and the essential conditions observed; but, for waves of increasing length, i. e. of smaller wave-number, the distances of the bands apart increases.

It should be noticed that the less the distance between the two sources, so much the greater is the distance apart of the bands. Further, since D, a, and c may all be measured, this experiment gives a method for the determination of the wave-length of the waves. It is not a very accurate

method, however; and so it is not often used. The method now universally adopted for the determination of wave-lengths will be explained later. (See Art. 365.)

363. Colors of Thin Plates. Another phenomenon due to interference is the so-called "colors of thin plates," such as are observed in soap-bubbles or any thin transparent films. It is noticed that when "white light" is incident upon such a transparent film, it appears colored both in transmitted and in reflected light. The film may be a solid, a liquid, or a gas. The explanation is not difficult. When waves of any definite wave-number fall upon a thin transparent film with parallel faces, some of the waves are reflected at the surface; others are refracted into the film. Of these waves refracted into the film, some are reflected at the lower surface, return to the upper surface and emerge parallel to those waves which were reflected on incidence; others, of course, are refracted out at the lower surface; and, of those reflected back to the upper surface, all do not emerge, but some may be reflected again, etc. It is evident that the waves coming away from the upper surface consist of several trains of waves all with their wave-normals parallel and all of the same wave-number; but the separate trains have passed over different paths and have been differently modified. Consequently, it may be that the superimposed waves are relatively retarded; the crest of one train may be behind the crest of another. This retardation depends upon three things: —

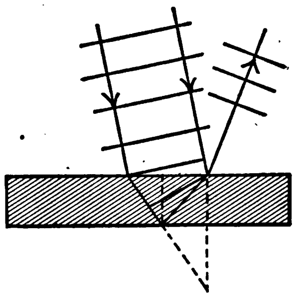


FIG. 275.

1. The actual difference in the paths of the different waves.

2. The different velocities of the waves in the film and outside.

3. The fact that one train of waves has been reflected when passing from a "faster" medium into a "slower," and the other while passing just the opposite way. This introduces a difference which is equivalent to $\frac{\lambda}{2}$ in the appearance of the reflected waves. (See Art. 135.)

If the total retardation of one train of waves behind the other is an odd number of half wave-lengths, the two waves will tend to neutralize each other. It may be shown that the intensities of the two interfering trains of waves are the same; and, consequently, under these conditions there will be absolutely no effect in the direction of the angle of reflection.

For different angles of incidence the retardation will naturally be different; and so, if the incidence of a train of waves of definite period upon a thin film is varied, there will be angles for which there will be no reflected waves, and others for which there will be. The retardation is different also for plates of different thickness. The transmitted waves are always such as will, together with the reflected waves, be equivalent to the incident waves; because there is supposed to be no absorption.

If, then, white light is incident upon a thin film at any angle, there will be some waves reflected (at the angle of reflection, of course); but corresponding to this angle of incidence and thickness of film, some particular train of waves will completely disappear, as just explained, from the reflected waves; and therefore the color-sensation produced by the reflected waves will be complementary to that which corresponds to the train which has disappeared. (Those waves which do not appear in the reflected waves are refracted out at the lower side; that is, are transmitted.) If the angle of incidence or the thickness of the film is varied, different trains of waves will be cut out; and so the color of the reflected light will change.

If a thick plate is used, instead of a thin one, several

different waves may be cut out from the incident waves; and those which are left may combine to produce practically a white sensation; so only thin plates are colored by the "white light."

There is a modification of this experiment which is called "Newton's Rings." A film of air is made by placing a convex lens upon a plane surface. There are thus circular strips, as it were, of varying thicknesses; and so, if a homogeneous train of waves falls upon this film, there will be a succession of thicknesses of these strips, which will not allow the waves to be reflected. Consequently there will be dark rings separated by bright ones. If white light is used, there is, of course, a general overlapping of the rings due to different waves, and certain beautifully colored rings appear.

In every case, of course, more of the waves pass through the film than are reflected, because at any transparent surface reflection is always weak. The transmitted effect, then, is always complicated by the presence of some waves which have come directly through. This is the reason why the transmission colors are always so much feebler than the reflected ones.

364. Diffraction by an Edge. In Chapter I., Article 303, it was explained why the illumination near the edge of a shadow is complicated by the hiding of portions of the Huyghens' zones by the obstacle which casts the shadow. This phenomenon is called "diffraction."

As the simplest case, consider waves of a definite wave-number coming from a slit S; and, after passing an opaque obstacle at R with its edge parallel to the slit, let them fall upon a screen placed at $P\ T\ Q$. The illumination at different points on the screen may be studied by constructing the wave-front at the instant when it reaches the obstacle, i.e. when its section is $O\ R\ O'$, and by drawing Huyghens' zones for each of the points. Thus, the pole of P is O, where $S\ O\ P$ is a radius of the wave-front; and

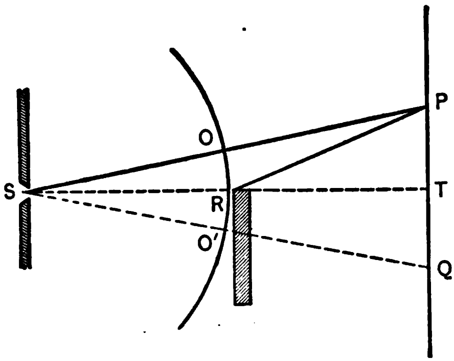


FIG. 276.

Huyghens' zones may be constructed around O, as explained in Chapter I. If the pole O is near the edge of the obstacle, that is, if the point P on the screen is near the edge of the shadow, it may not be possible to construct a large number of zones; and so the series $m_1 - m_2 + m_3 - m_4 +$ etc. does not have an infinite number of terms. If O is so near R that there is only a small number of complete zones, their effects at P will almost completely neutralize each other, if there is an *even* number of zones; while, if there is an *odd* number, the effect at P will be practically that due to the first zone. Therefore, for *small* distances on the wave-front away from R, the edge of the obstacle, the poles will be such that, as they recede from the edge, an even number of zones can be drawn, then an odd, then an even, and so on; and there will, therefore, be on the screen alternations in intensity, or bands, parallel to the edge of the obstacle and the slit. These bands are, of course, only near the edge of the shadow, just outside it in fact.

If Huyghens' zones are constructed for any point, Q, on the screen, behind the obstacle, the first zone which produces an effect will not be around its pole, O', but at the edge of the obstacle; and it follows that the entire effect at Q is due to one-half the first zone which it has at the edge R. If Q is far away from R, the effect is less than if

it is near; and, consequently, the illumination decreases uniformly for points below the edge of the obstacle.

Thus light waves do actually bend in slightly behind obstacles; but to a minute degree, owing to the shortness of the lengths of the waves.

These diffraction effects are observed near the edges of any shadows which are cast by small sources of light; for, if the source is large, the phenomenon is obscured by overlapping.

365. Diffraction-Grating. The "diffraction-grating" consists, in its simplest form, of a plate of glass on which are scratched by means of a diamond-point a series of parallel fine lines very close together and at exactly regular intervals. These scratches are, of course, opaque to light-waves, because they scatter and diffuse them; but in between them there are parallel transparent portions. As many as 15,000 or 20,000 lines in a distance of an inch (2.54 cm.) are often made. This ruling of a grating is done by means of an automatic machine, because everything must be perfectly regular.

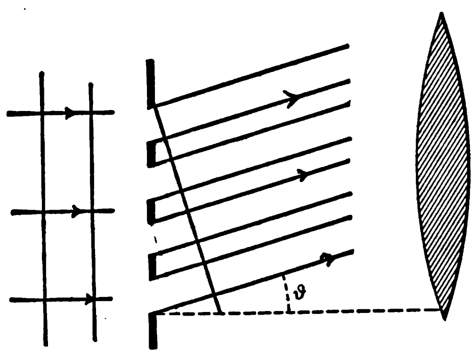


FIG. 277.

If ordinary "white light" is allowed to fall upon such a grating, the emerging light appears colored; and different colors will be seen on looking at the grating from different directions. The simplest mode of explanation is to show

how, if homogeneous light is incident on the grating, it is only at certain angles that any effect is transmitted.

The grating consists in principle of a series of transparent slits placed parallel to each other at regular intervals. A section of it would then be as shown. When a homogeneous train of plane waves reaches the grating, each opening becomes, as it were, a new source of light, and spherical waves spread out from each point. Consider the effect in any direction given by a line making an angle θ with the normal to the grating. All disturbances in this direction will be brought to focus in the focal plane of a converging lens, if it is placed so as to intercept the waves; and thus the actual effect may be studied. The distance from the edge of one opening to the corresponding edge of the other, a, is called the "grating-space;" and the difference in path between disturbances going from corresponding points in consecutive openings off in the direction θ is evidently $a \sin \theta$. Therefore, if the waves are falling perpendicularly upon the grating, so that the secondary waves from the grating-openings leave at the same instant, the effect from any one opening reinforces that from the next one, if

$$a \sin \theta = n \lambda,$$

where n is any whole number and λ is the wave-length of the waves; and so, if θ satisfies this condition, there will be a maximum effect produced at the corresponding point in the focal plane of the interposed lens. There will thus be points of maximum effect in the focal plane, corresponding to successive values of n; and, of course, the effect will be symmetrical on the two sides of the axis, for θ can be drawn below the axis as well as above it.

The general appearance on the focal plane may be represented as in the figure, where the elevations mean maximum effects or illumination. If there is a great number of lines in the grating, say 50,000 or 100,000,

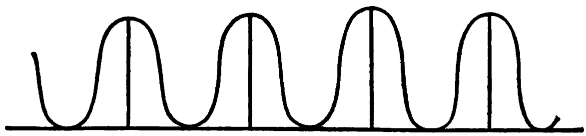


FIG. 278.

these curves for maximum illumination contract until, instead of there being a wide illumination around each position given by θ, there is only practically a straight line (if the waves come from a slit originally). Under these conditions it is possible to measure θ accurately. Thus, in the formula

$$a \sin \theta = n \lambda \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

a, the grating-space, may be measured; so can θ; and n, the "order" of the spectrum, is some known whole number, 1, 2, 3, etc. Consequently λ, the wave-length of the waves, may be measured most accurately.

In the accompanying table are given some values of the wave-lengths of certain absorption lines in the solar spectrum, which have received definite names.

TABLE XVII

WAVE-LENGTHS IN CENTIMETRES

K . . .	0.00003933825	D_2 . . .	0.00005890186
H . . .	0.00003968625	D_1 . . .	0.00005896357
g . . .	0.00004226904	C . . .	0.00006563054
G . . .	0.00004340634	B . . .	0.00006870186
F . . .	0.00004861527		
b_1 . . .	0.00005183791		
E_2 . . .	0.00005269723		
E_1 . . .	0.00005270495		

The unit commonly adopted for expression of wave-lengths is not the centimetre; but 0.00000001 of a centimetre.

Such a unit is called an Ångström unit, from the name of the great Swedish spectroscopist who first suggested its use. Thus, the wave-length of D_1 is 5896.357 Ångström units.

Since, then, for any value of λ there is a succession of definite values of θ for which there will be maximum effects; if white light is used, the different trains of waves will be spread out separately, and colored effects will be produced. It is also evident that a grating may take the place of a prism for the examination of the spectra of different sources; and so grating-spectroscopes are often used.

It should further be noted that the deviation and dispersion produced by a grating has not the faintest connection with its material; they are the same for all gratings of the same number of lines per inch.

Such a grating as has been just described is called a "transmission" one; because the waves are transmitted through its openings. Reflecting gratings may also be used, where an opaque mirror, e. g. polished metal, has a series of parallel scratches upon it at regular intervals. Again, instead of having the rulings on a plane surface, they may be on a curved surface; and concave reflecting gratings have certain most interesting properties, which cannot, however, be discussed here.

CHAPTER VII

DOUBLE REFRACTION

366. As explained in Chapter III., when a train of waves in any medium falls upon a surface which separates that medium from another, some waves enter the second medium, and have a definite wave-front. This fact is true of all isotropic media; but it is observed that, when waves pass into certain crystals in certain directions, there are two refracted wave-fronts. Such bodies as this are called "doubly refracting."

If plane waves fall upon a plate of a doubly refracting substance, which has parallel sides, there will be, in general, two trains of plane waves produced in the plate, in slightly different directions; and these two waves will emerge with their normals in the same direction as the incident waves. Following the path of any one disturbance, it is seen to produce two "rays." If a black point is placed on one side of the plate, there will be two virtual images of it at different points in the plate produced by the two sets of waves; and so, if the plate is viewed from the opposite side, two points will be apparently seen. (Compare the similar problem in a singly refracting plate, Art. 324.)

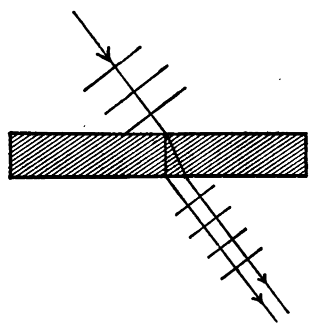


FIG. 279.

It is found by experiment that all crystals, with the exception of those which belong to the cubical system, are doubly refracting; and that isotropic substances become so if they are subjected to unequal strains, e. g. glass unevenly annealed. In certain crystals there is one direction in which waves may pass through without suffering double refraction; such are called "uniaxial." In other crystals there are two such directions; and such are called "biaxial."

367. Uniaxial Crystals. The best known examples of these are quartz and Iceland spar; and both are easily obtained and studied. It is found by experiment that of the two waves produced in general in a uniaxial crystal by a single incident train of waves, one obeys the laws of ordinary refraction, the other does not. The first train of waves is called "the ordinary wave;" the second, "the extraordinary;" and the latter, as a rule, obeys neither of the two laws of refraction. If a plate of Iceland spar is placed over a black point, two points are seen, as explained above; but, if now the plate is turned around in its own plane, the image due to the extraordinary waves will revolve around that due to the ordinary waves, which, of course, remains at rest.

368. Biaxial Crystals. In biaxial crystals neither train of waves obeys the ordinary laws of refraction in general, although for particular sections of the crystal one or both may. So in these crystals both trains of waves are extraordinary.

369. Since in a doubly refracting crystal the two trains of waves have different wave-points, it is evident that they travel with different velocities, and so have different indices of refraction. (Read Art. 319.) Therefore, if a prism is made of a doubly refracting substance, and if homogeneous plane waves fall upon its face, two plane waves will emerge, in different directions. If a converging lens is interposed, there will be two images in the focal plane

corresponding to the two directions of the waves. Thus the two trains of waves may be separated, and studied.

370. Nicol's Prism. In Iceland spar it may be proved that the index of refraction of the ordinary waves is greater than that of the extraordinary. A most interesting application of this fact has been made so as to completely separate the extraordinary and ordinary waves. In speaking of refraction (Art. 320), it was shown that if waves were passing from a medium of greater refractive index to one of less (i. e. of less velocity to greater), total reflection resulted if the angle of incidence exceeded a certain critical angle. It has been discovered that the index of refraction of Canada balsam is intermediate between those of the extraordinary and ordinary waves. A piece of Iceland spar may then be cut in two and gummed together again with Canada balsam; so that when, as a result of a single train of incident waves, the extraordinary and the ordinary waves fall upon the surface of the Canada balsam, the angle of incidence is such that it exceeds the critical angle for the ordinary waves, and so they will be totally reflected. The extraordinary waves, on the contrary, pass directly through the transparent balsam, losing very little light by reflection, and emerge on the farther side of the piece of Iceland spar. Thus the two waves are completely separated. Ordinarily, the sides of the Iceland spar are blackened so as to absorb the ordinary waves when they strike them; and thus only extraordinary waves can pass through. Such an arrangement as described is called a "Nicol's prism;" or more often, a "Nicol."

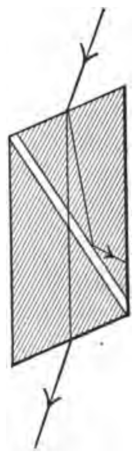


FIG. 280.

371. Property of Tourmaline. The uniaxial crystal tourmaline is doubly refracting; but, if its thickness exceeds more than 1 or 2 mm., only the extraordinary waves

emerge, as the ordinary ones are absorbed. Consequently, if ordinary light falls upon tourmaline, only one kind of waves emerge. Tourmaline is, however, a rather opaque crystal; and not much light is transmitted in any case. There are, however, other substances, more transparent, which have this same selective absorption.

CHAPTER VIII

POLARIZATION

So far, in the discussion of the phenomena associated with Light, explanations have been given in terms of the wave-theory; but nothing has been said as to the nature of the waves themselves. No phenomenon so far described requires the waves to be transverse, or longitudinal; but now certain phenomena will be discussed which do definitely require for their explanation that the ether-waves are transverse. These phenomena are generally classed under the name of "polarization," for reasons that will soon appear.

372. Polarization by Reflection. If a train of plane waves fall upon a plane piece of glass, or other transparent material, there will be both a reflected and a refracted train of waves; and, if this reflected train is made to fall upon a second plane piece of glass parallel to the first, there will also be a reflected and a refracted train in general. It is noticeable, though, that there are marked differences in the intensities of the reflected and the refracted trains at this second mirror; the latter being much the weaker. If the angle of incidence of the first train of waves on the first mirror is varied, an angle may be found for which there is absolutely no refracted train at all at the second glass mirror. If now the second mirror is revolved about an axis parallel to the wave-normal of the waves incident upon it, the angle of incidence will not change; but the plane of incidence, i. e. the plane including the normal to

the surface and the normal to the wave-front, will. When the plane of incidence has been thus revolved through 90°, absolutely no waves at all are reflected from the second mirror, but all the incident waves are refracted.

How is it possible to account for this fact, that waves reflected from one surface and falling upon another cannot be reflected? No explanation can be given if the ether-waves are longitudinal; for it is impossible to think of any modification in them which could prevent their reflection at the second surface. If the ether-waves are transverse, the explanation is not difficult. On this theory a train of ordinary ether-waves consists of the advance of some vibration which is perpendicular to the path of the waves. The vibration of any individual particle must be in the wave-front, but may be of any periodic nature; it may be in a straight line in any direction, it may be in a circle, it may be in an ellipse; and ordinary ether-waves may be regarded as a varying mixture of all these vibrations. But when such a train of waves reaches the surface of the glass at the definite angle of incidence, it may be supposed that only particular vibrations are reflected. There is almost conclusive evidence that with ordinary ether-waves only those vibrations, or components of vibrations, are reflected which are perpendicular to the plane of incidence, i. e. along the surface. (The other vibrations, i. e. those which are in the plane of incidence, and so strike down, as it were, into the surface, are refracted.) The plane surface of glass, then, for this angle of incidence allows only those vibrations to be reflected which are parallel to the surface. All the vibrations in the reflected wave are in the same direction; viz., parallel to a line perpendicular to the plane of incidence. Such a train of waves is said to be "plane polarized in the plane of incidence." When this train of waves falls upon the second glass mirror placed parallel to the first, its angle of incidence and plane of incidence are the same as they were for the first mirror; and therefore

all the waves will be reflected, the vibrations being parallel to the surface.

But, if the second mirror is so turned that the plane of incidence is at right angles to that for the first mirror, the vibrations in the incident waves are in the plane of incidence, and so enter the surface; consequently, since for this angle of incidence the surface reflects only those vibrations which are parallel to the surface, there is no reflected train of waves, all the incident waves are refracted.

This particular angle of incidence for which a plane transparent substance will reflect only vibrations parallel to its surface is called the "angle of polarization." In most cases, at the polarizing angle only a *maximum* amount of the waves whose vibrations are along the surface is reflected; that is, the polarization is not complete. But for those substances whose indices of refraction are in the neighborhood of 1.46, the polarization is complete at the polarizing angle. It is found by experiment that this angle of incidence is such that the reflected and refracted wave-normals are at right angles to each other. If a is the angle of polarization, it follows that the angle of refraction corresponding to it is $90° - a$. Hence if μ is the index of refraction,

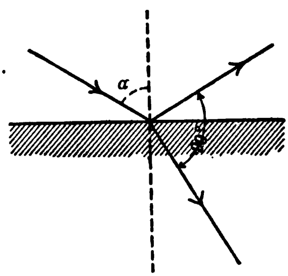


Fig. 281.

$$\frac{\sin a}{\sin (90° - a)} = \mu,$$

or

$$\tan a = \mu$$

Even if the angle of incidence is not the polarizing angle, there is always a proportion of the incident waves plane polarized by the reflection.

373. Polarization by Double Refraction. If the two trains

of plane waves produced in a doubly refracting substance are allowed in turn to fall upon a transparent mirror at its polarizing angle, it may be proved that they are both plane polarized, but that they are polarized in planes at right angles to each other. That is, if the mirror is so placed as to completely reflect one train of waves, the other will be completely refracted; and, if the mirror is turned around an axis parallel to the wave-normal of the incident waves until its plane of incidence has revolved 90°, the waves which were refracted before are now reflected and *vice versa*. This proves, then, that the vibrations in the two trains of waves are in straight lines at right angles to each other.

This fact furnishes at once a most convenient method for producing plane polarized waves. For a Nicol's prism or a piece of tourmaline allows only one train of waves to pass through; and consequently in the emerging waves the vibrations are all in one direction. There is, then, in a Nicol a certain fixed direction along which vibrations must be in order to be transmitted.

If waves emerging from one Nicol enter a second, they can emerge if the direction of the vibrations in the incident waves is that direction which the second Nicol can transmit, or if the vibration can have a component in that direction. If the Nicols are so placed that the two fixed directions are parallel, all the incident waves will be transmitted, and the Nicols are said to be "parallel." If, on the other hand, the two fixed directions are at right angles to each other, no waves will be transmitted, and the Nicols are said to be "crossed."

It should be observed that, since it is the extraordinary wave which is transmitted by a Nicol, the vibrations in this wave must be parallel to the fixed direction in the Nicol; the vibrations in the ordinary wave are at right angles to this direction. Consequently, when two Nicols are crossed, the vibrations which emerge from the first as

an extraordinary wave become the ordinary wave of the second, and are totally reflected at the layer of Canada balsam.

This fact also explains immediately an experiment of Huyghens', which consisted in looking at any small source of light through two plates of Iceland spar cut exactly alike and placed over each other. In general *four* images are seen; but, as one plate is turned around an axis perpendicular to itself, there are four positions 90° apart for which only two images are seen. The vibrations coming through the first plate are at right angles to each other, e. g. along a and b; but now, falling upon the second plate, in which there are only two directions which vibrations can have, e. g. a' and b', the vibration along a will have components along a' and b', and so will the vibration along b. Consequently there will be four emerging waves in general. (The vibrations emerge in two directions only; but, as the velocities of the two kinds of vibrations in the two plates are different, there will in general be four wave-fronts.) But if the two sets of perpendicular lines are parallel, there will obviously be only two emerging waves; and, as one plate is turned over the other, it will happen four times in one complete revolution that these lines are parallel.

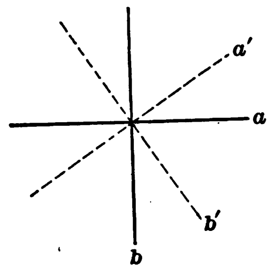


FIG. 282.

If a plane polarized wave falls upon a thin plate of a doubly refracting substance, its vibrations will be resolved into two at right angles to each other, because a doubly refracting substance allows in its transmitted waves only two directions of vibrations, which are at right angles, and which correspond to the two waves which characterize double refraction. It has been explained in the last chapter how double refraction is due to the fact that the two

sets of waves produced by it travel with different velocities. Consequently, when the plane polarized wave is resolved by the doubly refracting plate into two waves whose vibrations are at right angles to each other, these two waves will travel through the plate with different velocities; and one will emerge in advance of the other. The actual vibration in the emerging wave will be then a combination of two vibrations in straight lines at right angles to each other, one vibration being in general a trifle farther advanced than the other, i. e. one vibration may be at the end of its path when the other is just passing through its central position, etc.

374. Circularly Polarized Waves. If the thickness of the doubly refracting plate is such as to make one vibration come through in a time faster than the other by one quarter of a period, the vibration in the emerging wave is due to a combination of one vibration which is at the end of its path, x, when the other is just passing through the middle of its path, o, towards y. The actual path, or vibration, will, then, be an ellipse whose axes are the lines which mark the direction and extent of the component vibrations. The emerging waves are said to be "elliptically" polarized.

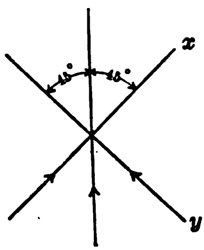


FIG. 283.

If, however, these two axes are of equal length, the vibration will be a circle; and the waves are said to be "circularly" polarized. If the two axes are of the same length, this means that the two component waves are of equal intensities; and this condition may be easily secured. For, if the direction of vibration of the incident plane polarized waves makes an angle of 45° with either one of the fixed directions of vibration in the doubly refracting plate, the incident vibrations will be resolved into two *equal* components along those two directions.

A plate of such a thickness as to produce a difference of time of a quarter of a period is called a "quarter-wave plate." So, in order to produce circularly polarized waves, it is simply necessary to allow ordinary plane waves to pass through a Nicol and then through a quarter-wave plate, the direction of vibration in the Nicol making an angle of 45° with the directions of possible vibration in the plate.

Conversely, if circularly polarized waves are passed through a quarter-wave plate, they become plane polarized, because the plate makes one of the component vibrations into which the circular vibration may be resolved lag behind the other quarter of a period; and two vibrations at right angles to each other, which pass through o together, or which reach x and y at the same instant, will combine to produce a vibration in a straight line. If a train of waves is, then, circularly polarized, the fact may be proved by seeing whether it is possible so to turn a Nicol that it prevents any waves passing after the train of waves has traversed a quarter-wave plate.

375. Interference of Polarized Waves. Two waves cannot completely interfere unless their vibrations are in the same direction; and, if any one of the interference experiments in Article 362 is modified by placing a Nicol's prism in front of each of the two sources of light so that the two interfering waves are plane polarized, it may be proved that interference-bands do not occur when the two Nicols are crossed. If they are parallel, the bands have a maximum intensity; while in intermediate positions the bands are feebler.

This experiment proves conclusively that ether-waves are transverse.

376. Colors Due to Polarization. When homogeneous waves are passed through a thin plate of a doubly refracting substance, there are two emerging sets of vibrations, at right angles to each other; and one set is retarded be-

hind the other, as explained in Article 374. These two waves cannot interfere, because the vibrations are at right angles to each other; but, if these waves now pass through a Nicol's prism, they will each produce a train of waves whose vibrations are in the same direction, viz., the fixed direction of the Nicol, the initial vibrations being resolved into components in this direction and at right angles to it, the latter not emerging. Consequently, the two emerging waves are now in a condition to interfere. If the waves which are incident on the doubly refracting plate are ordinary waves, the mode of vibration in them will be constantly changing; and so these two waves which emerge from the Nicol will not have any constant character. If, however, the incident waves are plane polarized by being passed through a Nicol, the two emerging waves will have definite intensities. One of the waves will be retarded in the doubly refracting plate behind the other; and it may be easily proved that, for definite thicknesses of the plate, the two sections of the homogeneous waves will completely interfere. For waves of different wave-numbers different thicknesses of the plate will serve to absorb them. Consequently, if "white light" passes through a Nicol, then through a thin doubly refracting plate, and finally enters a second Nicol, a certain definite color will be absent; and so its complementary color will be seen. (This is true except for four positions of the two Nicols.) If the doubly refracting plate is of varying thickness, there will be a definite color corresponding to each thickness. The colors observed when the Nicols are in any one position are complementary to those observed when the position of one Nicol is turned through 90°.

If the Nicols are crossed, no light at all would come through them if the doubly refracting plate were not interposed between them. This fact of the appearance of color-phenomena when a thin plate of doubly refracting substance is placed between two crossed Nicols is the best

test there is for the detection of double refraction in any substance.

There are many other ways of producing colored phenomena by means of polarized light-waves, but they are too complicated for discussion here.

377. Rotation of the Plane of Polarization. As already noted in ELECTRICITY (Art. 295), when a plane polarized wave is passed along the lines of force in a magnetic field, the plane of polarization is rotated; that is, the direction of vibration changes.

Many natural substances produce rotation of the plane of polarization also, — quartz, if cut with its faces perpendicular to its axis; and many "active" organic chemicals, such as some tartaric acids and some sugars in aqueous solution. In certain cases the plane is rotated towards the right, in others towards the left; but in all these substances, if the plane polarized waves are made to retrace their paths by being sent back in the opposite direction, exactly the opposite change takes place; the rotation, as it were, untwists.

This last fact is not true of the rotation produced in a magnetic field; because there, entirely independently of whether the waves are in the same or the opposite direction along the lines of force, the rotation is always connected with the direction of the lines of force by the "right-handed screw law."

It should be noted that in both cases the amount of the rotation is different for waves of different wave-numbers, being least for waves of the least wave-numbers.

INDEX

Numbers refer to Pages